МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ

ХАРКІВСЬКА НАЦІОНАЛЬНА АКАДЕМІЯ МІСЬКОГО ГОСПОДАРСТВА

МЕТОДИЧНІ ВКАЗІВКИ
до самостійної роботи та практичних занять
з курсу

ТЕОРІЯ ІМОВІРНОСТЕЙ ТА МАТЕМАТИЧНА СТАТИСТИКА

(для слухачів другої вищої освіти ФПО та ЗН спеціальностей
7.050106 – "Облік і аудит",
7.050107 – "Економіка підприємства")

Харків
ХНАМГ
2012

Методичні вказівки до самостійної роботи та практичних занять з курсу «Теорія імовірностей та математична статистика» (для слухачів другої вищої освіти ФПО та ЗН спеціальностей 7.050107 - "Економіка підприємства", 7.050106 - "Облік і аудит") / Харк. нац. акад. міськ. госп-ва; уклад.: В. М. Охріменко, Т. Б. Воронкова, О. О. Воронков. - Х. : ХНАМГ, 2012. - 68 с.

Укладачі : доц. В. М. Охріменко,
ст. викл. Т. Б. Воронкова,
ст. викл. О. О. Воронков

Рекомендовано кафедрою „Економіки підприємств міського господарства”, протокол № 01 від 30.08.11 р.

ЗАГАЛЬНІ ПОЛОЖЕННЯ

Курс «Теорія імовірностей і математична статистика» є нормативною дисципліною у навчальному плані напряму «Економіка і підприємництво». Обсяг курсу становить 108 академічних годин або 3 кредити ECTS. Програма курсу розділена на три змістових модуля: «Теорія імовірностей», «Математична статистика» та «Випадкові процеси».

Метою вивчення дисципліни «Теорія імовірностей і математична статистика» є формування базових знань в області застосування імовірнісно-статистичного апарата, вивчення закономірностей у масових випадкових явищах, визначення їх імовірнісних характеристик з метою прогнозування.

В результаті вивчення курсу студенти повинні оволодіти основними методами визначення імовірнісних характеристик випадкових величин, статистичного опису результатів спостереження та перевірки статистичних гіпотез для прийняття на їх основі обґрунтованих рішень. Методичні вказівки спрямовані на допомогу студентам оволодіти практичними навичками із застосування імовірнісно-статистичного апарата.

«Теорія імовірностей і математична статистика» вивчає закономірності у випадкових явищах та є основою для побудови кількісних моделей керування економічними системами. Прикладами таких моделей є моделі планування та керування запасами, теорії ігор, теорії масового обслуговування. Імовірнісно-статистичні методи є базовими для теорії прийняття рішень - складової частини сучасного менеджменту. Статистичні показники аналізують при оцінці ризику в інвестиційній діяльності, у діяльності страхових компаній, а також у багатьох сферах економіки та управління.

Широка сфера застосування теорії імовірностей і математичної статистики зумовлює важливе місце, що займає даний курс у підготовці економістів вищої кваліфікації. Ця дисципліна закладає теоретичну основу для наступного вивчення курсів «Математичне програмування», «Економетрія», «Статистика», «Економічний аналіз», «Економічні ризики», «Теорія прийняття рішень».

МЕТОДИЧНІ ВКАЗІВКИ ДО САМОСТІЙНОЇ РОБОТИ

ЗМ 1. ТЕОРІЯ ІМОВІРНОСТЕЙ

Тема 1. ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ІМОВІРНОСТЕЙ

Основні поняття та визначення теорії імовірностей: випадкові й детерміновані явища, дослід, випадкова подія, імовірність. Класичний та статистичний методи визначення імовірності випадкової події. Поняття про частоту випадкової події.

Література: [1] с. 7-16; [3] с. 12-16; [4] с. 15-34.

Запитання для самоперевірки:

1. Дайте визначення випадкової події.
2. Які події називаються: а) достовірними? б) рівноможливими? в) несумісними? г) протилежними? Наведіть приклади.
3. Чи є протилежні події несумісними?
4. Чи є несумісні події протилежними?
5. Дайте визначення імовірності випадкової події.
6. Як підрахувати імовірність події класичним методом?
7. Що розуміють під повною групою подій? Наведіть приклади.
8. Чи завжди можна визначити імовірність випадкової події класичним методом?
9. Як пов'язані між собою імовірність і частота появи події?

Задачі для самостійного розв'язання:

1.1. На складі 20 деталей, з яких 17 придатних. Визначити імовірність того, що з трьох навмання взятих деталей всі виявляться придатними.

1.2. На складі 50 придатних і 5 дефектних деталей. Визначити імовірність того, що серед п'яти навмання взятих деталей одна виявиться дефектною.

1.3. Учасники жеребкування тягнуть із ящика жетони з номерами від 1 до 100. Знайти імовірність того, що номер першого жетона, який навмання витягнутий з ящика, не містить цифру 5.

1.4. У партії з 20 готових виробів є 4 бракованих. Партію ділять на дві рівні частини. Визначити імовірність того, що браковані вироби розділяться нарівно.

1.5. Набираючи номер телефону, абонент забув останні дві цифри і, пам'ятаючи, що вони різні, набрав їх навмання. Знайти імовірність того, що були набрані потрібні дві цифри.

1.6. У розіграші першості з баскетболу беруть участь 18 команд, з яких випадковим способом формуються дві групи по 9 команд у кожній. Серед учасників змагань є 5 команд екстракласу. Знайти імовірність того, що а) всі команди екстракласу потраплять в одну групу; б) дві команди потраплять в одну з груп, а три - в іншу.

1.7. Є дві урни, у першій з яких a білих та b чорних кулі, у другій – c білих та d чорних. З кожної урни виймають по одній кулі. Знайти імовірність того, що обидві вийнятих кулі опиняться білими.

1.8. У ліфт будинку, в якому сім поверхів, на першому поверсі ввійшли три пасажири. Кожний з них з однаковою імовірністю виходить на кожному з поверхів. Знайти імовірність того, що всі пасажири вийдуть одночасно (на тому самому поверсу).

Тема 2. ОПЕРАЦІЇ НАД ПОДІЯМИ. ТЕОРЕМИ ТЕОРІЇ ІМОВІРНОСТЕЙ. ОСНОВНІ ФОРМУЛИ ТЕОРІЇ ІМОВІРНОСТЕЙ

Сума та добуток випадкових подій. Теорема додавання для несумісних та сумісних подій. Протилежні події. Теорема множення для залежних та незалежних подій. Залежні події. Умовна імовірність випадкової події. Формула повної імовірності. Формула Бейєса (теорема гіпотез). Повторні незалежні випробування. Локальна та інтегральна теореми Лапласа. Формула Пуассона.

Література: [1] с. 17-20; 23-28; 31-51; [3] с. 17-33; [4] с. 37-81.

Запитання для самоперевірки:

1. Як визначити імовірність суми сумісних подій?
2. Чи може сума двох подій збігатися з їхнім добутком?
3. Наведіть приклади залежних і незалежних подій.
4. Що розуміють під умовною імовірністю події?
5. Як визначається імовірність добутку двох подій?
6. В яких випадках для визначення імовірності застосовується формула Бернуллі?
7. Дайте визначення найімовірнішого числа появ події.
8. Як обчислити найімовірніше число появ події?
9. Чим розрізняються задачі, в яких потрібне застосування локальної і інтегральної граничних теорем?
10. В яких випадках замість формули Бернуллі використовується формула Пуассона?

Задачі для самостійного розв'язання:

2.1. На кожній з 5 однакових карток надруковано одну з наступних букв: «О», «П», «Р», «С», «Т». Картки ретельно перемішані. Знайти імовірність того, що на вийнятих по одній та розташованих в одну лінію картках з'явиться слово «СПОРТ».

2.2. В урні a білих та b чорних куль. З урни одну за одною виймають дві кулі. Знайти імовірність того, що обидві кулі будуть білі.

2.3. В урні a білих та b чорних куль. З урни одну за одною виймають всі кулі, що знаходяться у ній. Знайти імовірність того, що другою буде вийнято білу кулю.

2.4. Два стрілки незалежно один від одного роблять два постріли (кожний у свою мішень). Імовірність влучення у мішень при одному пострілі для першого стрільця p_1, для другого p_2. Виигравшим змагання вважається той стрілець, у мішені якого буде більше пробоїн. Знайти імовірність того, що виграє перший стрілець.

2.5. Імовірність того, що протягом однієї зміни виникне неполадка верстата, дорівнює 0,05. Визначити імовірність того, що протягом трьох змін верстат не зламається жодного разу.

2.6. Прилади одного найменування виготовляють два заводи. На виробництво, що використовує прилади, 2/3 приладів надходять з першого заводу та 1/3 - з другого. Імовірність безвідмовної роботи приладу (надійність), виготовленого першим заводом, дорівнює 0,95, а другим - 0,9. Визначити повну надійність приладу, що надійшов на виробництво.

2.7. Службовці підприємства розподілені за підрозділами та статтю в такий спосіб: у виробничому відділі - 8 жінок і 18 чоловіків, у плановому відділі - 4 жінки і 9 чоловіків, у відділі реалізації - 10 жінок і 5 чоловіків. Навмання обраний службовець виявився чоловіком. Визначити імовірність того, що він працівник а) виробничого відділу; б) планового відділу; в) відділу реалізації.

2.8. Готові вироби містять 5% браку. Визначити імовірність того, що в числі п'яти взятих навмання виробів: а) немає жодного дефектного; б) два дефектних.

2.9. Виробництво дає 1% браку. Визначити імовірність того, що з 1500 виробів бракованих буде не більше 20.

Тема 3. ВИПАДКОВА ВЕЛИЧИНА ТА ЇЇ ЗАКОНИ РОЗПОДІЛУ

Поняття випадкової величини. Безперервні й дискретні випадкові величини. Закон розподілу випадкової величини. Ряд розподілу. Функція розподілу та її властивості. Щільність розподілу та її властивості.

Література: [1] с. 57-59; 105-117; [3] с. 34-39; [4] с. 82-99.

Запитання для самоперевірки:

1. Дайте визначення випадкової величини.
2. Яка випадкова величина називається дискретною? Наведіть приклади.
3. Яка випадкова величина називається безперервною? Наведіть приклади.
4. Поясніть, з якою метою в теорії імовірностей розрізняють дискретні і безперервні випадкові величини?
5. Що має на увазі термін «закон розподілу»? В яких формах може бути представлений закон розподілу випадкової величини?
6. Чи може функція розподілу бути: а) більше одиниці; б) від’ємною?
7. Що розуміють під щільністю розподілу випадкової величини?
8. Чому не має сенсу поняття щільності розподілу для дискретної випадкової величини?
9. Яка розмірність щільності розподілу?
10. Перелічіть властивості щільності розподілу.
11. Як, маючи ряд розподілу, знайти значення функції розподілу?
12. Як виражається імовірність влучення випадкової величини на інтервал значень, якщо відомо функцію розподілу? Щільність розподілу?

Задачі для самостійного розв'язання:

3.1. Розглядаючи невипадкову величину C як окремий вид випадкової, побудувати для неї функцію розподілу, знайти математичне сподівання і дисперсію.

3.2. У партії з 30 виробів є 7 дефектних. З цієї партії випадковим способом обрані три вироби для перевірки їхньої якості. Побудувати ряд розподілу числа відібраних для перевірки виробів (випадкової величини X).

3.3. Для умов попередньої задачі побудувати функцію розподілу числа відібраних для перевірки виробів (випадкової величини X).

3.4. Випадкова величина X задана функцією розподілу

$$F(x)=\begin{cases}0, & x<2\\ (x-2)^2, & 2\le x\le 3\\ 1, & x>3\end{cases}$$

Знайти щільність розподілу f(x) і визначити імовірність влучення випадкової величини X в інтервал (1,5; 2).

Тема 4. ЧИСЛОВІ ХАРАКТЕРИСТИКИ ВИПАДКОВОЇ ВЕЛИЧИНИ

Середнє значення та математичне сподівання. Мода. Медіана. Моменти випадкової величини: початкові й центральні. Дисперсія та середнє квадратичне відхилення. Теореми про числові характеристики.

Література: [1] с. 67-70; 77-84; [3] с. 40-45; [4] с. 107-121.

Запитання для самоперевірки:

1. Назвіть основні числові характеристики випадкових величин.
2. Як пов'язані між собою математичне сподівання і середнє арифметичне значень випадкової величини?
3. Математичне сподівання - випадкова величина чи ні?
4. Чи є дисперсія випадковою величиною?
5. Як математичне сподівання і дисперсія характеризують випадкову величину?
6. Чим зручне застосування замість дисперсії середнього квадратичного відхилення?
7. В яких одиницях вимірюють математичне сподівання?
8. В яких одиницях вимірюють дисперсію?
9. Чому дорівнює математичне сподівання невипадкової величини C?
10. Як мода і медіана характеризують випадкову величину?

Задачі для самостійного розв'язання:

4.1. Для умов прикладу 3.2 визначити математичне сподівання, дисперсію і середнє квадратичне відхилення випадкової величини.

4.2. Для умов прикладу 3.4 визначити математичне сподівання, дисперсію і середнє квадратичне відхилення випадкової величини.

4.3. До випадкової величини X додали невипадкову величину *a*. Як від цього зміняться її числові характеристики: математичне сподівання, дисперсія і середнє квадратичне відхилення?

4.4. Випадкову величину X помножили на невипадкову величину a. Як від цього зміняться її числові характеристики: математичне сподівання, дисперсія і середнє квадратичне відхилення?

4.5. Виконують два незалежних постріли по мішені. Імовірність влучення при кожному пострілі дорівнює р. Розглядаються дві випадкові величини: X- різниця між числом влучень і числом промахів і Y - сума числа влучень і числа промахів. Побудувати для випадкових величин X і Y ряд розподілу (для кожної окремо) і знайти їхні числові характеристики.

Тема 5. НАЙВАЖЛИВІШІ ДЛЯ ПРАКТИКИ ЗАКОНИ РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН

Біноміальний закон розподілу. Закон розподілу Пуассона. Експонентний закон розподілу. Поняття найпростішого потоку подій. Число подій, що потрапляють на ділянку часу τ. Проміжок часу між двома сусідніми подіями в найпростішому потоці T. Нормальний закон розподілу імовірностей. Інтеграл імовірностей. Правило трьох сигма. Поняття про центральну граничну теорему. Закон рівномірної щільності.

Література: [1] с. 60-67; 118-131; 146-150; [3] с. 46-61; [4] с. 129-150; 153-167.

Запитання для самоперевірки:

1. Яким умовам повинні задовольняти повторні незалежні випробування?
2. Як визначають числові характеристики випадкової величини, розподіленої за законом Бернуллі?
3. Який зв'язок існує між біноміальним і пуассонівським розподілами?
4. Яким умовам повинна задовольняти випадкова величина, підпорядкована закону Пуассона?
5. Як визначають числові характеристики закону розподілу Пуассона?
6. Якими параметрами визначається експонентний закон розподілу випадкової величини?
7. Чому дорівнює щільність імовірності випадкової величини з нормальним законом розподілу?
8. Якими параметрами визначається нормальний закон розподілу випадкової величини?
9. Як змінюється графік нормального закону із зміною середнього квадратичного відхилення випадкової величини?
10. Як визначити імовірність влучення нормально розподіленої випадкової величини на задану ділянку?
11. Поясніть імовірнісний смисл параметрів нормального розподілу.
12. Поясніть смисл центральної граничної теореми.

Задачі для самостійного розв'язання:

5.1. Випадкова величина X підлегла закону розподілу Пуассона з математичним сподіванням a=3. Побудувати многокутник розподілу і функцію розподілу випадкової величини X. Знайти імовірність того, що випадкова величина X прийме значення менше, ніж її математичне сподівання.

5.2. Випадкова величина X підпорядкована експонентному закону розподілу з параметром μ:

$$f(x)=\begin{cases}\mu e^{-\mu x} & при\ x>0\\ 0 & при\ x<0\end{cases}$$

Побудувати криву розподілу, визначити функцію розподілу і знайти імовірність того, що випадкова величина X прийме значення, менше її математичного сподівання.

5.3. Вимірювальний прибор має систематичну помилку 5 м і середню квадратичну помилку 75 см. Яка імовірність того, що помилка вимірювання не перевершить за абсолютною величиною 5 м?

5.4. Випадкова величина X підпорядкована нормальному закону з математичним сподіванням, рівним нулю. Імовірність влучення цієї випадкової величини на ділянку від $-\alpha$ до $+\alpha$ дорівнює 0,5. Знайти середнє квадратичне відхилення і написати вираз нормального закону.

5.5. У світлофорі на перехресті 1 хвилину горить зелене світло та 0,5 хвилини - червоне. Автомобіль під'їжджає до перехрестя у випадковий момент, не пов'язаний з роботою світлофора. Знайти імовірність того, що він проїде перехрестя, не зупиняючись.

Тема 6. СИСТЕМА ВИПАДКОВИХ ВЕЛИЧИН. ЗАКОНИ РОЗПОДІЛУ ТА ЧИСЛОВІ ХАРАКТЕРИСТИКИ СИСТЕМИ

Багатовимірна випадкова величина. Поняття системи випадкових величин. Система двох випадкових величин. Функція розподілу та щільність розподілу імовірностей системи двох випадкових величин, їх властивості. Числові характеристики системи, кореляційний момент та коефіцієнт кореляції. Функції випадкових величин.

Література: [1] с. 135-142; 153-164; 177-181; [3] с. 62-72; [4] с. 177-190; 213-219; 258-276.

Запитання для самоперевірки:

1. Що являє собою багатомірна випадкова величина?
2. Що являє собою функція розподілу системи двох випадкових величин? Перелічіть її властивості.
3. Перелічіть числові характеристики системи двох випадкових величин.
4. Що характеризує кореляційний момент системи двох випадкових величин?
5. Для чого використовується коефіцієнт кореляції?
6. Перелічіть теореми про числові характеристики.
7. Чому дорівнює середнє квадратичне відхилення добутку невипадкової величини C на випадкову величину X?
8. Сформулюйте теорему додавання математичних сподівань для випад-

кових величин: а) залежних і незалежних; б) корельованих і некорельованих.

9. Чому дорівнює математичне сподівання добутку двох незалежних випадкових величин?

Задачі для самостійного розв'язання:

6.1. Два стрілки незалежно один від іншого проводять по одному пострілу, кожний по своїй мішені. Випадкова величина X – число влучень першого стрільця, Y– число влучень другого стрільця. Імовірність влучення в мішень для першого стрільця дорівнює p_1=0,9, а для другого p_2=0,8. Побудувати функцію розподілу системи випадкових величин X і Y.

6.2. Для умов попереднього прикладу визначити числові характеристики випадкового вектора (X, Y).

6.3. Незалежні випадкові величини X і Y розподілені за нормальними законами з параметрами m_x=2; m_y=-3; σ_x=1; σ_y=2. Визначити імовірність події $A=\{X< m_x$ і $Y<m_y\}$.

6.4. Відомі математичне сподівання і дисперсія випадкової величини X: m_x=2; D_x=3. Знайти математичне сподівання і дисперсію випадкової величини Y=3X-2.

6.5. Випадкові величини X і Y мають математичні сподівання m_x=-1, m_y=1 і дисперсії D_x=4 і D_y=9. Знайти математичне сподівання випадкової величини Z=3XY + 5.

Тема 7. ЗАКОН ВЕЛИКИХ ЧИСЕЛ

Принцип практичної впевненості. Формулювання закону великих чисел. Рівень значущості. Лема Маркова. Нерівність Чебишева. Теорема Чебишева. Теорема Бернуллі.

Література: [1] с. 94-104; [3] с. 73-79; [4] с. 399-413; [5] с. 180-196.

Запитання для самоперевірки:

1. Що називається законом великих чисел? Поясніть смисл цієї назви.
2. Яка роль закону великих чисел у теорії імовірностей?
3. У чому полягає принцип практичної впевненості?
4. Поясніть смисл поняття «рівень значущості».
5. Сформулюйте теорему Чебишева і поясніть, в чому полягає її практичний зміст.
6. Сформулюйте теорему Бернуллі і поясніть, в чому полягає її практичний зміст.
7. Чи можна стверджувати, що при нескінченно великій кількості дослідів n частота події p* дорівнює імовірності цієї події p? Обґрунтуйте відповідь.

Задачі для самостійного розв'язання:

7.1. Вага виробу, що виготовляється підприємством, є випадковою величиною з математичним сподіванням 90 г і дисперсією 0,0225. Визначити імовірність того, що відхилення ваги виробу від її середнього значення за абсолютною величиною не перевищить 0,4 г. Для розв'язання використати нерівність Чебишева.

7.2. З 1000 виробів, що надходять у складальний цех, випадковим способом вибрали 200 виробів для контролю. Серед них виявилося 25 бракованих. Прийнявши частку бракованих виробів з контрольної партії як імовірність виготовлення бракованого виробу, оцінити імовірність того, що у всій партії виробів, які надійшли у складальний цех, бракованих виявиться не менше 10 % і не більше 15 %. Для розв'язання використати теорему Бернуллі.

ЗМ 2. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ

Тема 8. ОБРОБКА СТАТИСТИЧНИХ ДАНИХ

Вибірковий метод. Поняття генеральної та вибіркової сукупностей. Варіанта. Частота. Варіаційний ряд. Полігон розподілу. Кумулятивна крива. Гістограма розподілу. Визначення закону розподілу спостережуваної ознаки за статистичними даними. Числові характеристики варіаційного ряду. Варіаційний розмах R. Середнє лінійне відхилення d. Дисперсія варіаційного ряду. Стандартне відхилення. Коефіцієнт варіації V. Властивості вибіркових числових характеристик. Довірчий інтервал та довірча імовірність.

Література: [1] с. 185-210; 214-227; [3] с. 80-107; [4] с. 430-444; 451-467; [5] с. 205-207; 217-223; 237-241; 311-316;.

Запитання для самоперевірки:

1. Які завдання вирішує математична статистика? Назвіть основні з них.
2. Поясніть зміст вибіркового методу.
3. У чому полягає різниця між генеральною сукупністю і вибіркою?
4. Яку інформацію про досліджувану ознаку дістають з варіаційного ряду?
5. Що таке оцінка параметра розподілу?
6. Якими властивостями повинні володіти вибіркові числові характеристики варіаційного ряду?
7. Поясніть властивості спроможності й незмещеності оцінок.
8. Чим відрізняються точкова і інтервальна оцінки параметрів розподілу?
9. Поясніть поняття «довірчий інтервал» і «довірча імовірність».

Задачі для самостійного розв'язання:

8.1. На біржі протягом деякого часу проводяться статистичні дослідження коливань ціни на партії товарів. Результати спостережень: 3100, 4000, 3800, 4100, 3400, 4200, 3700, 3900, 3200, 4100, 3800, 4200, 3500, 4000, 3900. Побудувати варіаційний ряд і гістограму випадкової величини X - ціна на партію товарів.

8.2. Для умов попередньої задачі обчислити числові характеристики варіаційного ряду.

8.3. Для умов задач 8.1 і 8.2 визначити довірчі інтервали для вибіркової середньої з надійністю 0,95 і вибіркової дисперсії з надійністю 0,99. Прийняти, що досліджувана ознака X розподілена нормально.

Тема 9. ЕЛЕМЕНТИ ТЕОРІЇ КОРЕЛЯЦІЇ

Функціональна, статистична та кореляційна залежності. Рівняння регресії. Поле кореляції. Коефіцієнт регресії. Метод найменших квадратів. Вибірковий коефіцієнт кореляції. Вибіркове кореляційне відношення. Міжгрупова та внутрігрупова дисперсії.

Література: [1] с. 249-260; 271-279; [3] с. 108-131; [4] с. 445-450; [5] с. 365-381;.

Запитання для самоперевірки:

1. Які задачі вирішують методом кореляційного аналізу?
2. В яких випадках залежність y = f(x) є функціональною, статистичною або кореляційною?
3. Дайте визначення термінів «регресія», «лінія регресії», «рівняння регресії».
4. Поясніть значення термінів «пояснююча змінна», «результативна ознака».
5. З яких міркувань визначають тип кореляційної залежності y= f(x)? Які типи залежностей Ви знаєте?
6. Чим характерна лінійна залежність y = f(x)? Чому її використовують найбільш часто?
7. Як називаються параметри лінійної залежності y = f(x)?
8. Які методи можна використовувати для визначення параметрів рівняння регресії y = f(x)?
9. Якій вимозі задовольняють параметри, визначені за методом найменших квадратів?
10. Назвіть характеристики, що дозволяють оцінити наявність зв'язку між ознакою-фактором і результативною ознакою.
11. Які значення може приймати коефіцієнт кореляції, які висновки можна зробити на підставі цих значень?
12. Які значення може приймати кореляційне відношення, і які висновки можна зробити на підставі цих значень?
13. Що таке кореляційна таблиця?
14. Які параметри визначають за допомогою кореляційної таблиці?

Задачі для самостійного розв'язання:

9.1. Результати вимірювань досліджуваної ознаки Y зведені в таблицю

x_i	41	50	81	104	120	139	154	180	208	241	250	269	301
y_i	4	8	10	14	16	20	14	23	26	30	31	36	37

Використовуючи поле кореляції, вибрати клас залежності y=((x), побудувати рівняння регресії Y на X, оцінити тісноту зв'язку між фактором і результативною ознакою.

9.2. Результати вимірювань фактору X і результативної ознаки Y наведені в таблиці

x_i	1	2	3	4	5	6	7	8
y_i	6	17	34	57	86	121	162	209

Користуючись методом найменших квадратів визначити параметри апроксимуючої залежності $y=ax^2+bx+c$.

9.3. У результаті статистичних спостережень зареєстрована залежність $u=f(t)$.

u_i	75	55	40	30	20	15	10	10	5	5
t_i	1	2	3	4	5	6	7	8	9	10

Визначити параметри експонентної апроксимуючої залежності.

Тема 10. ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ

Поняття статистичної гіпотези. Нульова та альтернативна гіпотези. Критична область. Область прийняття гіпотези. Критична точка zk. Статистичний критерій. Помилки 1-го та 2-го роду. Потужність критерію. Порівняння вибіркової середньої та генеральної середньої нормальної сукупності. t-критерій (розподіл Стьюдента). Порівняння двох дисперсій нормальних генеральних сукупностей. F-критерій (розподіл Фішера). Порівняння вибіркової та генеральної дисперсій нормальної сукупності. Критерій $\chi 2$. Перевірка гіпотези про нормальний розподіл генеральної сукупності. Елементи дисперсійного аналізу. Загальна, факторна та залишкова сума квадратів відхилень. Загальна, факторна та залишкова дисперсії. Елементи регресійного аналізу.

Література: [1] с. 282-290; 332-338; 342-351; [3] с. 132-162; [4] с. 445-450; 467-470; [5] с. 281-287.

Запитання для самоперевірки:

1. Поясніть, що таке статистична гіпотеза?
2. Поясніть, що таке нульова і альтернативна гіпотези?
3. Які критерії застосовують для перевірки статистичних гіпотез?
4. Поясніть значення термінів «критична область», «критична точка», «область прийняття гіпотези».
5. В якому випадку слід вибирати двосторонню критичну область?
6. Який результат перевірки гіпотези відносять до помилки 1-го роду та який до помилки 2-го роду?
7. Наведіть приклади задач на перевірку гіпотез.
8. Які задачі вирішують за допомогою однофакторного дисперсійного аналізу?

9. Чому цей вид аналізу випадкових даних отримав назву дисперсійного?

10. Що являють собою величини $S_{общ}$, $S_{факт}$ і $S_{ост}$? Яке співвідношення між ними?

11. Які задачі вирішують методом регресійного аналізу?

12. Для чого застосовують метод найменших квадратів?

13. Що таке рівняння регресії? Як перевіряють його адекватність статистичним даним?

14. Як визначають значущість коефіцієнтів рівняння регресії?

Задачі для самостійного розв'язання:

10.1. Число появ герба при 20 підкиданнях двох монет розподілилося в такий спосіб:

Число гербів	0	1	2
Число появ	4	8	8

Чи узгоджуються ці дані з припущенням про симетричність монет і незалежність результатів підкидання. Як рівень значущості прийняти α=0,05.

ЗМ 3. ВИПАДКОВІ ПРОЦЕСИ

Тема 11. ЕЛЕМЕНТИ ТЕОРІЇ ВИПАДКОВИХ ПРОЦЕСІВ

Поняття випадкового процесу. Імовірнісні характеристики випадкового процесу. Кореляційна функція та її властивості. Лінійне перетворення суми випадкових функцій Xi(t). Стаціонарний випадковий процес. Математичне сподівання та дисперсія стаціонарного випадкового процесу. Кореляційна функція стаціонарного випадкового процесу. Властивість ергодичності. Числові характеристики, визначені за множиною реалізацій та за часом. Елементи теорії масового обслуговування. Основні поняття. Класифікація СМО. Вхідний потік вимог. Дисципліна черги. Механізм обслуговування. Поняття марківського випадкового процесу. Випадковий процес з дискретними станами та безперервним часом. Граф станів системи. Рівняння для імовірностей станів. Одноканальна СМО з відмовами. Фінальні імовірності станів. Характеристики системи масового обслуговування. Багатоканальна СМО з відмовами. Одноканальна СМО з очікуванням та обмеженою чергою. Одноканальна СМО з очікуванням та необмеженою чергою. Багатоканальна СМО з очікуванням та необмеженою чергою. Комп’ютерна обробка статистичних даних.

Література: [3] с. 163-203; [5] с. 386-409; 412-421.

Запитання для самоперевірки:

1. Дайте визначення випадкового процесу.
2. Що таке реалізація випадкової функції.
3. Які властивості імовірнісних характеристик стаціонарного випадкового процесу?
4. Дайте визначення кореляційної функції.

5. Поясніть властивість ергодичності стаціонарного випадкового процесу.

6. Які задачі вирішує теорія масового обслуговування?

7. Перелічіть основні компоненти системи масового обслуговування, що обумовлюють її математичний опис.

8. Поясніть класифікацію систем масового обслуговування.

9. В якому випадку випадковий процес, що протікає в системі, називається марковським?

10. Дайте визначення стаціонарного потоку подій і регулярного потоку подій.

11. Які змінні є невідомими в рівняннях Колмогорова? Що таке фінальної імовірності?

12. Якими параметрами визначається потік заявок? Потік обслуговувань?

13. В чому полягає суть методу статистичних випробувань?

14. В чому полягає суть моделювання у нейронних мережах?

Задачі для самостійного розв'язання:

9.1. Випадкова функція X(t) в кожному перетині являє собою безперервну випадкову величину з щільністю розподілу f(x,t). Напишіть вираз для математичного сподівання $m_x(t)$ і дисперсії $D_x(t)$ випадкової функції X(t).

9.2. Випадкова функція X(t) задана у вигляді X(t) = Vt+1, де V – випадкова величина, розподілена за нормальним законом з параметрами m_v=1 і σ_v=0,2. Визначити числові характеристики випадкової функції X(t).

9.3. Розглядається робота автоматичної телефонної станції, розрахованої на одночасне обслуговування 20 абонентів. Виклик на АТС надходить в середньому через 6 секунд. Кожна розмова триває в середньому 2 хвилини. Якщо абонент застає АТС зайнятою, він отримує відмову. Якщо абонент застає вільним хоча б один з 20 каналів, він з'єднується з потрібним йому номером. Визначити імовірність того, що абонент буде обслужений, а також інші характеристики роботи СМО: середнє число зайнятих каналів, імовірність зайнятості каналу, середній час простою каналу.

МЕТОДИЧНІ ВКАЗІВКИ ДО ПРАКТИЧНИХ ЗАНЯТЬ

Практичне заняття 1

ВИЗНАЧЕННЯ ІМОВІРНОСТІ ВИПАДКОВОЇ ПОДІЇ

Мета — сформувати вміння визначати імовірності випадкових подій з використанням теорем теорії імовірностей.

У теорії імовірностей поняття випадкової події є одним з основних. *Випадкова подія* - це будь-який факт, що в наслідку досліду може відбутися або не відбутися.

Імовірністю випадкової події називають числову міру ступеню об'єктивної можливості появи цієї події в наслідку досліду.

Імовірність випадкової події можна визначити за класичним методом, якщо наслідки досліду *утворюють повну групу, є рівноможливими* та *несумісними*. Імовірність події А визначається як відношення числа можливих наслідків досліду, яки сприяють появі події А, до загального числа можливих наслідків досліду:

$$P(A) = \frac{m}{n}$$

де n - загальне число можливих наслідків досліду; m - число наслідків досліду, яки сприяють появі події А.

Сумою двох подій А і В називають подію С, яка полягає у появі події А або події В або обох подій разом: С = А + В.

Добутком двох подій А і В називають подію С, що полягає у спільній появі подій А і В: С = А * В.

Протилежними називають дві несумісних події А і $\bar{A}$, якщо вони складають повну групу.

Подію А називають *незалежною* від події В, якщо імовірність події А не зміниться від того, відбулася подія В чи ні. Якщо ж імовірність події А залежить від того, відбулася подія В чи ні, то такі події називають *залежними*.

Імовірність події А, обчислену за умови, що подія В відбулася, називають *умовною імовірністю* події А.

Теорема додавання. Імовірність суми двох несумісних подій А і В дорівнює сумі імовірностей цих подій, тобто

$$P(A + B) = P(A) + P(B)$$

Слідства теореми додавання:

1. Якщо події A_1, A_2, … , A_i, …, A_n утворюють повну групу несумісних подій, сума їх імовірностей дорівнює 1:

$$P(\Sigma A_i) = \Sigma P(A_i) = 1.$$

2. Сума імовірностей двох протилежних подій дорівнює одиниці:

$$P(A) + P((A) = 1, \text{ або}$$

$$P(A) = 1 - P(\overline{A}).$$

Якщо дві події є сумісними, імовірність їх суми дорівнює сумі імовірностей цих подій мінус імовірність їх спільної появи:

$$P(A + B) = P(A) + P(B) - P(A*B).$$

Імовірність суми n сумісних подій

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_i P(A_i) - \sum_{i,j} P(A_i A_j) + \sum_{i,j,k} P(A_i A_j A_k) - \ldots + (-1)^{n-1} P(A_1 A_2 \ldots A_n),$$

де суми поширюються на всі можливі комбінації індексів i, j, k,..., узятих по одному, по два, по три і т.д.

Теорема множення. Імовірність добутку двох подій A і B дорівнює добутку імовірності одного з них на умовну імовірність іншого, обчислену за умови, що перша відбулася:

$$P(A * B) = P(A) * P(B|A).$$

Для імовірності добутку n подій формула має вигляд

$$P(A_1 A_2 \ldots A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) \ldots P(A_n | A_1 A_2 \ldots A_{n-1})$$

Якщо події A і B незалежні, то умовна імовірність події B дорівнює безумовній імовірності цієї події,

$$P(B|A) = P(B).$$

Імовірність добутку двох незалежних подій дорівнює добутку імовірностей цих подій:

$$P(A * B) = P(A) * P(B).$$

Якщо маємо n незалежних подій:

$$P\left(\prod_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i).$$

Формула повної імовірності. Повна безумовна імовірність події A з урахуванням випадковості умов протікання досліду, дорівнює сумі добутків імовірностей кожної з гіпотез на умовну імовірність події A при кожній з гіпотез.

$$P(A) = \sum_{i=1}^{N} P(H_i A) = \sum_{i=1}^{N} P(H_i) P(A / H_i).$$

Формула Бейеса (теорема гіпотез). Ця формула дозволяє переоцінити імовірності гіпотез після того, як стає відомим наслідок досліду, в результаті якого відбулася подія A. За відомими до проведення досліду (апріорними) імовірностями гіпотез $P(H_i)$ та за результатом досліду (настання події A) обчислюють післядослідні (апостеріорні) імовірності гіпотез $P(H_i|A)$.

$$P(H_i / A) = \frac{P(H_i) * P(A / H_i)}{\sum P(H_i) * P(A / H_i)}.$$

Формула Бернуллі (повторні незалежні випробування). Імовірність того, що в результаті певного числа дослідів подія A з'явиться рівно m разів

$$P_n(m) = C_n^m p^m q^{n-m}.$$

де $P_n(m)$ - імовірність того, що в n випробуваннях подія A з'явиться рівно m разів; C_n^m - число сполучень з n елементів по m; p – імовірність появи події A в одному досліді; q = 1 – p – імовірність не появи події A в одному досліді.

Локальна теорема Лапласа: Якщо імовірність p появи події А в кожному випробуванні постійна і відмінна від нуля та одиниці, імовірність $P_n(m)$ того, що подія А з'явиться в n дослідах рівно m разів, приблизно дорівнює, і тім точніше, чим більше n, значенню функції

$$P_n(m)=\frac{1}{\sqrt{npq}}\varphi(x),$$

де $x=\dfrac{m-np}{\sqrt{npq}}$, а значення функції $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-x^2/2}$ визначаються за довідковими таблицями. Функція ((x) парна, тобто φ(-x) = φ(x).

Інтегральна теорема Лапласа: Якщо імовірність p настання події А в кожному випробуванні постійна і відмінна від нуля та одиниці, приблизно імовірність $P_n(m_1,m_2)$, того, що подія А з'явиться у випробуваннях від m_1 до m_2 разів,

$$P_n(m_1,m_2)=\frac{1}{\sqrt{2\pi}}\int_{x'}^{x''}e^{-x^2/2}dx$$

де $x'=\dfrac{m_1-np}{\sqrt{npq}}$; $x''=\dfrac{m_2-np}{\sqrt{npq}}$.

Формула Пуассона. Якщо число незалежних випробувань n велике, але значення добутку np залишається невеликим, імовірність того, що в цих випробуваннях подія А відбудеться m разів, можна визначити за формулою:

$$P_n(m)=\frac{(np)^m}{m!}e^{-np}.$$

Вказівки до виконання завдання

Після уважного вивчення умови завдання треба сформулювати подію, імовірність якої треба визначити. Потім обґрунтувати, які теореми теорії імовірностей (теорему додавання або теорему множення) або формули (формулу повної імовірності, формулу Бейєса та ін.) треба застосувати для розв’язання задачі. У процесі розв’язання задачі треба чітко висловлювати міркування.

Задача 1.1

Партія виробів містить N виробів, з яких M є дефектними. Навмання з цієї партії вибирають k виробів для контролю. Визначити імовірність того, що серед них буде рівно *l* дефектних виробів.

Розв'язання

Запишемо подію, для якої необхідно визначити імовірність

A={у контрольній партії рівно *l* дефектних виробів}.

Загальне число можливих наслідків досліду дорівнює $n=C_N^k$. Число наслідків досліду, які сприяють появі події А={у контрольній партії рівно *l* дефектних виробів} визначимо в такий спосіб. Число випадків, які сприяють появі *l* дефектних виробів

$$m_д = C_m^l \ .$$

Число випадків, які сприяють появі k-l не дефектних виробів

$$m_г = C_{N-M}^{k-l} \ .$$

Імовірність події А

$$P(A) = \frac{m_д * m_г}{n} = \frac{C_m^l * C_{N-M}^{k-l}}{C_N^k} \ .$$

Задача 1.2

Студент прийшов здавати іспит. Він знає 15 з 20 питань програми. Визначити імовірність того, що він відповість на три пропонованих екзаменаційних питання.

Розв'язання

Запишемо подію, імовірність якої необхідно визначити, А={студент знає відповіді на три запитання}. Виразимо її через елементарні події:

A_1 ={знає відповідь на перше запитання}

A_2 ={знає відповідь на друге запитання}

A_3 ={знає відповідь на третє запитання}

$$A = A_1 * A_2 * A_3.$$

Події A_1, A_2, A_3 – залежні події. Обчислимо їх умовні імовірності.

$$P(A_1) = \frac{15}{20}.$$

Умовна імовірність події A_2, за умови, що відбулася подія A_1,

$$P(A_2|A_1) = \frac{14}{19}.$$

Умовна імовірність події A_3, за умови, що відбулися події A_1 і A_2,

$$P(A_3|A_1 * A_2) = \frac{13}{18}.$$

Тоді імовірність події А за теоремою множення:

$$P(A) = P(A_1) * P(A_2|A_1) * P(A_3|A_1 * A_2) = \frac{15}{20} * \frac{14}{19} * \frac{13}{18} = 0{,}368.$$

Задача 1.3

Ви повинні знайти людину, в якої день народження збігається з Вашим. Скількох незнайомців Вам треба буде опитати, щоб імовірність події А={день народження людини збігається з Вашим} була не меншою за 0,5?

Розв'язання

Імовірність того, що перша людина, в якої Ви запитали, не народилася в один день із Вами, дорівнює

$$P(\overline{A}) = \frac{365-1}{365}.$$

Якщо Ви опитаєте n чоловік, то за теоремою множення імовірність, що вони не народилися в один день із Вами дорівнюватиме:

$$P(\overline{A}) = (\frac{365-1}{365})^n.$$

Тоді імовірність події A={день народження людини збігається з Вашим}:

$$P(A) = 1 - P(\overline{A}) = 1 - (\frac{365-1}{365})^n.$$

Оскільки задана ймовірність P(A) = 0,5, маємо:

$$0{,}5 = 1 - \left(\frac{365-1}{365}\right)^n.$$

Звідки n = 253 чоловік.

Задача 1.4

У двох урнах знаходяться кулі, що відрізняються тільки кольором. У першій - 5 білі; 11 чорних та 8 червоних. У другій - 10 білих; 8 чорних та 6 червоних. З кожної урни виймають одну кулю. Визначити імовірність події A={вийняті кулі однакового кольору}.

Розв'язання

Для визначення імовірності події A ={вийняті кулі однакового кольору} виразимо її через елементарні події:

A_1 ={обидві кулі білі}

A_2 ={обидві кулі чорні}

A_3 ={обидві кулі червоні}.

Імовірність події A за теоремою множення для незалежних подій:

$$P(A) = P(A_1) * P(A_2) * P(A_3).$$

Тоді:

$$P(A) = 5/24 * 10/24 + 11/24 * 8/24 + 8/24 * 6/24 = 0{,}32.$$

Задача 1.5

Є три однакові на вигляд урни. У першій - 2 білі та 3 чорних кулі, в другій - 4 білі та 1 чорна, у третій - 3 білі. Навмання з однієї з урн виймають одну кулю. Визначити імовірність того, що вийнята куля виявиться білою.

Розв'язання

Позначимо подію A={вийнята куля біла}. Висуваємо три гіпотези:

H_1 ={обрана перша урна}

H_2 ={обрана друга урна}

H_3 ={обрана третя урна}

Імовірність кожної гіпотези

$$P(H_1) = P(H_2) = P(H_3) = 1/3.$$

Умовні імовірності події А:

$$P(A|H_1)=\frac{2}{5};\qquad P(A|H_2)=\frac{4}{5};\qquad P(A|H_3)=1.$$

Повна безумовна ймовірність події А

$$P(A)=P(H_1)*P(A|H_1)+P(H_2)*P(A|H_2)+P(H_3)*P(A|H_3)=$$
$$=\frac{1}{3}*\frac{2}{5}+\frac{1}{3}*\frac{4}{5}+\frac{1}{3}*1=\frac{11}{15}.$$

Задача 1.6

Два стрільці роблять по одному пострілу у мішень. Для першого з них імовірність влучення дорівнює 0,8, а для другого - 0,4. Мішень пробита один раз (одне влучення). Знайти імовірність того, що мішень уражена першим стрільцем.

Розв'язання

Є факт, тобто подія А={мішень уражена один раз}, тобто один із стрільців промахнувся. Висуваємо гіпотези: H_1={мішень уражена першим стрільцем}; H_2={мішень уражена другим стрільцем}. Визначимо імовірності гіпотез. Мішень уражена першим стрільцем, якщо він при пострілі потрапив у мішень, а другий стрілець промахнувся, тоді $P(H_1)=0{,}8*(1-0{,}4)=0{,}48$. Мішень уражена другим стрільцем, якщо він при пострілі потрапив у мішень, а перший стрілець промахнувся, тоді $P(H_2)=(1-0{,}8)*0{,}4=0{,}08$. Умовна імовірність події А, за умови, що має місце гіпотеза H_1 дорівнює $P(A|H_1)=1$, і за умови, що має місце гіпотеза H_2 дорівнює $P(A|H_1)=1$, тому що в цих випадках мішень буде напевно уражена один раз. Скористуємося теоремою гіпотез і визначимо імовірність реалізації гіпотези H_1:

$$P(H_i/A)=\frac{P(H_i)*P(A/H_i)}{\sum P(H_i)*P(A/H_i)}=\frac{0{,}48*1}{0{,}48*1+0{,}08*1}=0{,}884.$$

Задача 1.7

Є п'ять однотипних пристроїв. Імовірність безвідмовної роботи кожного дорівнює 0,8. Визначити імовірність того, що в робочому стані перебувають m пристроїв (m = 1, 2, 3, 4, 5).

Розв'язання

Визначимо імовірність, що всі п'ять пристроїв пошкоджені:

$$P_5(0)=C_5^0*0{,}8^0*0{,}2^{5-0}=0{,}00032;$$

визначимо імовірність, що чотири пристрої пошкоджені:

$$P_5(1)=C_5^1*0{,}8^1*0{,}2^{5-1}=0{,}00638;$$

визначимо імовірність, що три пристрої пошкоджені:

$$P_5(2)=C_5^2*0{,}8^2*0{,}2^{5-2}=0{,}0512;$$

визначимо імовірність, що два пристрої пошкоджені:

$$P_5(3)=C_5^3*0{,}8^3*0{,}2^{5-3}=0{,}2047;$$

визначимо імовірність, що один пристрій пошкоджений:

$$P_5(4) = C_5^4 * 0{,}8^4 * 0{,}2^{5-4} = 0{,}4095;$$

визначимо імовірність, що жодний пристрій не пошкоджений:

$$P_5(5) = C_5^5 * 0{,}8^5 * 0{,}2^{5-5} = 0{,}328.$$

$$\Sigma p_i = 0{,}00032 + 0{,}00638 + 0{,}0512 + 0{,}2047 + 0{,}4095 + 0{,}328 = 1.$$

Задача 1.8

Імовірність того, що виписаний продавцем чек буде оплаченим, дорівнює 0,9. Яке найімовірніше число чеків буде оплачене, якщо виписано 40 чеків?

Розв'язання

Отже, $n = 40$, $p=0{,}9$, $q=1-0{,}9=0{,}1$.

1) Скористаємося для розв'язання формулою $np - q \le m_0 \le np + p$:

$$40*0{,}9-0{,}1 \le m_0 \le 40*0{,}9 + 0{,}9$$

$$35{,}9 \le m_0 \le 36{,}9.$$

Цієї нерівності задовольняє ціле число $m_0 = 36$.

2) Обчислимо добуток $np = 40 * 0{,}9 = 36$. Оскільки він є цілим числом, $m_0 = 36$.

Задача 1.9

Оптова база постачає товари у 10 магазинів, від кожного з яких може надійти заявка на черговий день з імовірністю 0,4, незалежно від заявок інших магазинів. Знайти найімовірніше число заявок у день та імовірність одержання цього числа заявок.

Розв'язання

У цьому випадку $n = 10$, $p=0{,}4$. Добуток np дорівнює цілому числу $np = 10*0{,}4 = 4$. Імовірність того, що від магазинів буде отримано 4 заявки, можна визначити за формулою Бернуллі:

$$P_n(m) = C_n^m p^m q^{n-m},$$

де $m=4$, $q=1-0{,}4=0{,}6$. Тоді імовірність того, що від магазинів буде отримано 4 заявки:

$$P_{10}(4) = C_{10}^4 0{,}4^4 0{,}6^{10-4} = 0{,}251.$$

Задача 1.10

Визначимо імовірність того, що деяка подія А в 150 дослідах з'явиться рівно 12 разів. Імовірність появи події А в одному досліді дорівнює 0,1.

Розв'язання

У цьому випадку $n = 150$, $p=0{,}1$. Тоді $q=1-0{,}1=0{,}9$, $m=12$. Для визначення імовірності, скористаємося локальною теоремою Лапласа:

$$P_n(m)=\frac{1}{\sqrt{npq}}\varphi(x).$$

Обчислимо $x=\frac{m - np}{\sqrt{npq}}=\frac{12-150*0,1}{\sqrt{150*0,1*0,9}}=-0,82$, значення функції визначимо за довідковою таблицею φ(-0,82) =0,2939, тоді:

$$P_{150}(12)=\frac{1}{\sqrt{150*0,1*0,9}}*0,2939=0,08.$$

Визначимо імовірність того, що подія А в 150 дослідах з'явиться не менше за 10 і не більше за 20 разів. Для цього скористаємося інтегральною теоремою Лапласа.

Знайдемо $x'=\frac{m_1-np}{\sqrt{npq}}=\frac{10-150*0,1}{\sqrt{150*0,1*0,9}}=-1,36$ та $x''=\frac{m_2-np}{\sqrt{npq}}=\frac{14-150*0,1}{\sqrt{150*0,1*0,9}}=-0,27$.

$$P_n(m_1,m_2) = Ф(x'')-Ф(x') = 0,4131 – 0,1064 = 0,307.$$

Задача 1.11

Імовірність виготовлення нестандартної деталі дорівнює 0,004. Знайти ймовірність того, що серед 1000 деталей буде 5 нестандартних.

Розв'язання

Формалізуємо задачу: n = 1000, p = 0,004, *a* = np = 1000 * 0,004 = 4. Для знаходження ймовірності події $P_{1000}(5)$ уживемо формулу Пуассона:

$$P_{1000}(5)=\frac{(4)^5}{5!}e^{-4}=0,1563.$$

Практичне заняття 2
ЗАКОНИ РОЗПОДІЛУ ТА ЧИСЛОВІ ХАРАКТЕРИСТИКИ ВИПАДКОВОЇ ВЕЛИЧИНИ

Мета — сформувати вміння на підставі універсального закону розподілу випадкової величини визначати імовірності її значень та числові характеристики.

Випадковою називають величину, яка в результаті досліду може прийняти те або інше значення. *Дискретною* називають випадкову величину, число значень якої скінченне, або нескінченне, але рахункове (яка може приймати тільки окремі значення). *Безперервною* називають випадкову величину, число значень якої нескінченне навіть на невеликому інтервалі.

Законом розподілу випадкової величини називають будь-яке правило, що дозволяє будь-якому значенню випадкової величини поставити у відповідність його імовірність.

Ряд розподілу - це таблиця, у верхньому рядку якої перелічені всі значення випадкової величини $x_1, x_2, \ldots, x_n$ в порядку зростання, а у нижньому - імовірності появи цих значень $p_1, p_2, \ldots, p_n$:

x_i	x_1	x_2	…	x_n
p_i	p_1	p_2	…	p_n

де $p_i = P\{X=x_i\}$.

Оскільки події $\{X=x_1\}$, $\{X=x_2\},\ldots,\{X=x_n\}$несумісні та утворюють повну групу, сума їх імовірностей дорівнює одиниці $\sum p_i =1$.

Найбільш загальною формою закону розподілу для всіх випадкових величин (дискретних та безперервних) є функція розподілу.

Функція розподілу випадкової величини X - це імовірність того, що випадкова величина X прийме значення, менше за x:

$$F(x) = P\{X \le x\}.$$

Функція розподілу має наступні властивості.

1. Значення функції розподілу належать відрізкові [0; 1]: $0\le F(x_2) \le 1$;

2. Функція розподілу - неубутна функція, тобто $F(x_2) \ge F(x_1)$, якщо $x_2 > x_1$.

3. Імовірність того, що випадкова величина X прийме значення, укладене в інтервалі (x_1, x_2), дорівнює приросту функції розподілу на цьому інтервалі:

$$P\{ x_1 \le X \le x_2\}= F(x_2) - F(x_1).$$

4. Якщо всі можливі значення випадкової величини X належать інтервалу $(-\infty, +\infty)$, то при мінус нескінченності функція розподілу дорівнює нулю, а при плюс нескінченності - одиниці, тобто $F(-\infty) = 0$; $F(+\infty) = 1$.

Щільністю розподілу випадкової величини X у точці x називається похідна функції розподілу X у цій точці (передбачається, що F(x) безперервна і диференційована):

$$f(x) = \frac{dF(x)}{dx}.$$

Властивості щільності розподілу імовірностей: 1. Щільність розподілу є невід'ємною, тобто $f(x) \ge 0$ як похідна неубутної функції; 2. Функція розподілу визначається за співвідношенням:

$$F(x) = \int_{-\infty}^{x} f(x)dx.$$

3. Інтеграл від щільності розподілу у нескінченних межах дорівнює одиниці:

$$\int_{-\infty}^{+\infty} f(x)dx = 1,$$

де f(x)dx - елемент імовірності, тобто імовірність влучення випадкової величини X на елементарну ділянку dx.

4. Імовірність влучення безперервної випадкової величини на інтервал (x_1, x_2) дорівнює інтегралу щільності розподілу в межах від x_1 до x_2.

$$P\{x_1 \le X \le x_2\} = \int_{x_1}^{x_2} f(x)dx.$$

Математичним сподіванням випадкової величини X називають суму добутків всіх можливих її значень на імовірності цих значень

$$M[X] = \sum_{i=1}^{n} x_i p_i .$$

Для безперервної випадкової величини математичне сподівання

$$m_x = \int_{-\infty}^{+\infty} xf(x)dx .$$

Математичне сподівання характеризує середнє значення випадкової величини.

Другий початковий момент α_2:

$$\alpha_2 = M[X^2]$$

Під центрованою випадковою величиною розуміють її відхилення від математичного сподівання:

$$\overset{0}{X} = X - m_x .$$

Дисперсія випадкової величини для дискретної X:

$$D_x = \sum_{i=1}^{n} \overset{0}{x}_i{}^2 p_i = \sum_{i=1}^{n} (x_i - m_x)^2 p_i .$$

Для безперервної випадкової величини:

$$D_x = \int_{-\infty}^{+\infty} \overset{0}{x}{}^2 f(x)dx = \int_{-\infty}^{+\infty} (x - m_x)^2 f(x)dx .$$

Дисперсія випадкової величини є характеристикою розсіювання цієї величини навколо математичного сподівання.

Середнє квадратичне відхилення X:

$$\sigma_x = \sqrt{D_x}.$$

Вказівки до виконання завдання

Після уважного вивчення умови завдання треба сформулювати, імовірність якої події потребує визначити завдання. Потім обґрунтувати, якими співвідношеннями можна скористуватися для розв'язання задачі. У процесі розв'язання треба чітко висловлювати міркування та поняття теорії імовірностей.

Задача 2.1

Тричі кидають монету. Випадкова величина X - число появ герба. Побудувати ряд розподілу випадкової величини X. Визначити функцію розподілу випадкової величини X та побудувати її графік.

Розв'язання

Побудуємо ряд розподілу X. Очевидно, що число появ герба при триразовому киданні монети може приймати чотири значення 0, 1, 2, 3. Для визначення імовірностей цих значень скористаємося формулою Бернуллі. Число дослідів n=3, імовірність появи герба в одному досліді p=0,5, імовірність непояви герба в одному досліді q=1-p=1-0,5=0,5. Отже, ряд розподілу випадкової величини X має вигляд:

x_i	0	1	2	3
p_i	1/8	3/8	3/8	1/8

Виконаємо перевірку: $\sum p_i$ = 1/8+3/8+3/8+1/8=1.

За визначенням функція розподілу випадкової величини X - це імовірність того, що випадкова величина X прийме значення, менше x: $F(x) = P\{X \leq x\}$.

Обчислимо значення функції розподілу:

F(0) = P{X (0} = 0;

F(1) = P{X (1} = 1/8;

F(2) = P{X (2} = 1/8 + 3/8 = 4/8;

F(3) = P{X (3} = 4/8 + 3/8 = 7/8;

при X(3 F(x) = 1.

Побудуємо графік F(x) (рис. 2.1)

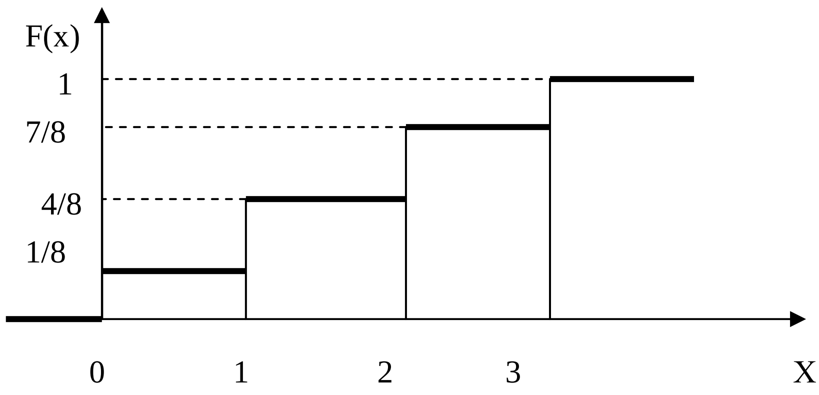


Рис. 2.1 - Графік функції розподілу

Задача 2.2

Випадкову величину X задано функцією розподілу:

$$F(x) = \begin{cases} 0, x \leq 0 \\ x^2 / 4, 0 < x \leq 2 \\ 1, x > 2 \end{cases}.$$

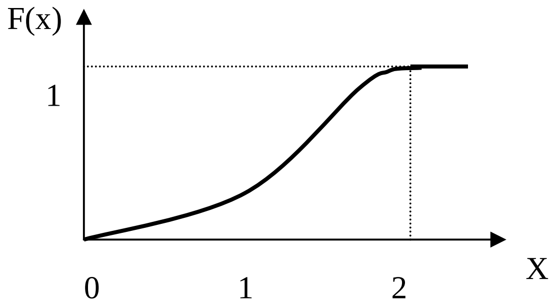
Знайти щільність розподілу імовірностей та побудувати графіки функцій f(x) та F(x).

Розв'язання

Знайдемо щільність імовірностей, взявши похідну від функції розподілу на кожному інтервалі:

$$f(x) = \begin{cases} 0, x \le 0 \\ x/2, 0 < x \le 2 \\ 0, x > 2 \end{cases}.$$

в) Побудуємо графіки функцій f(x) та F(x).



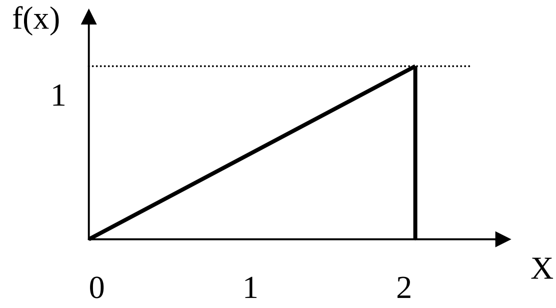


Рис. 2.2 - Графік функції розподілу та щільності імовірностей

Задача 2.3

Визначимо числові характеристики дискретної випадкової величини для умов прикладу задачі 2.1. Маємо ряд розподілу:

x_i	0	1	2	3
p_i	1/8	3/8	3/8	1/8

Розв'язання

Визначимо математичне сподівання випадкової величини X:

$$m_x = \Sigma x_i * p_i = 0*1/8 + 1*3/8 + 2*3/8 + 3*1/8 = 1{,}5.$$

Дисперсію визначимо за двома способами:

за формулою другого центрального моменту:

$D_x = \Sigma(x_i - m_x)^2 * p_i = (0-1{,}5)^2 * 1/8 + (1-1{,}5)^2 * 3/8 + (2-1{,}5)^2 * 3/8 + (3-1{,}5)^2 * 1/8 = 0{,}75.$

та за формулою, що містить другий початковий момент α_2:

$$D_x = \alpha_2 - m_x^2$$

$$\alpha_2 = \Sigma x_i^2 * p_i = 0^2 * 1/8 + 1^2 * 3/8 + 2^2 * 3/8 + 3^2 * 1/8 = 3.$$

$$D_x = \alpha_2 - m_x^2 = 3 - 1{,}5^2 = 0{,}75.$$

Визначимо середнє квадратичне відхилення випадкової величини X:

$$\sigma_x = \sqrt{D_x} = \sqrt{0{,}75} = 0{,}855.$$

Задача 2.4

Нехай випадкову величину X задано функцією розподілу:

$$F(x) = \begin{cases} 0, x \le 0 \\ x^2/4, 0 < x \le 2 \\ 1, x > 2 \end{cases}.$$

Визначити числові характеристики безперервної випадкової величини.

Розв'язання

Знайдемо щільність імовірностей:

$$f(x)=\begin{cases}0, x\le 0\\ x/2, 0<x\le 2\\ 0, x>2\end{cases}.$$

Визначимо математичне сподівання X:

$$M[X]=\int_0^2 x*x/2dx=1/2\int_0^2 x^2dx=\frac{1}{2}\frac{x^3}{3}\Big|_0^2=4/3.$$

Знайдемо дисперсію X:

$$D[X]=\int_0^2 x^2*x/2dx-(4/3)^2=1/2\int_0^2 x^3dx-(4/3)^2=\frac{1}{2}\frac{x^4}{4}\Big|_0^2-(4/3)^2=2/9.$$

Практичне заняття 3
НАЙВАЖЛИВІШІ ДЛЯ ПРАКТИКИ ЗАКОНИ РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН

Мета — сформувати вміння користуватися законами розподілу випадкових величин для визначення їх числових характеристик та імовірностей того, що випадкова величина належатиме певному інтервалу її значень.

Біноміальний закон розподілу. Дискретна випадкова величина X має біноміальний закон розподілу (розподіл Бернуллі), якщо її можливі значення: 0, 1, ..., n, а відповідні імовірності визначаються за співвідношенням:

$$P_n(m)=C_n^m p^m q^{n-m}.$$

де p - імовірність появи події A в одному досліді, $0<p<1$; q - імовірність не появи події A в одному досліді, $q=1-p$.

Числові характеристики випадкової величини, розподіленої за біноміальним законом мають вигляд:

$$m_x = np;\ D_x = npq;\ \sigma_x = \sqrt{npq}.$$

Закон розподілу Пуассона визначає імовірність того, що за якийсь час τ відбудеться рівно k подій за законом Пуассона визначається формулою:

$$P(k)=\frac{(\lambda\tau)^k}{k!}e^{-\lambda\tau},$$

де λ - число подій в одиницю часу; τ - інтервал часу.

Математичне сподівання (середнє число подій, що потрапляють на ділянку часу довжиною τ) та *дисперсія* випадкової величини визначаються формулою:

$$m_x = D_x = \lambda\tau.$$

Експонентний закон розподілу. Функцію розподілу T обчислюють за формулою: $F(t) = P\{T < t\} = 1 - e^{-\lambda t}$.

Щільність розподілу T як похідна функції розподілу F(t) має вигляд:

$$f(t) = d(t)/dt = \lambda e^{-\lambda t}.$$

Математичне сподівання випадкової величини, розподіленої за експонентним законом, зворотно параметру розподілу λ: $m_x = 1/\lambda$.

Дисперсія: $D_t = \dfrac{1}{\lambda^2}$, середнє квадратичне відхилення: $\sigma_x = 1/\lambda$.

Імовірність влучення випадкової величини, що має експонентний розподіл, в інтервал значень (α, β): $P\{\alpha \le t \le \beta\} = e^{-\lambda\alpha} - e^{-\lambda\beta}$.

Нормальний закон розподілу імовірностей визначається двома параметрами m_x і σ_x.

Імовірність влучення випадкової величини X на ділянку значень (α, β) виражається через функцію Лапласа формулою:

$$P\{\alpha < X < \beta\} = \Phi\left(\frac{\beta - m_x}{\sigma}\right) - \Phi\left(\frac{\alpha - m_x}{\sigma}\right).$$

Функція розподілу для випадкової величини, розподіленої нормально визначається за формулою

$$F(x) = \Phi\left(\frac{x - m_x}{\sigma_x}\right) + 0{,}5 \cdot$$

Імовірність влучення нормально розподіленої випадкової величини в ділянку значень, симетричну щодо її математичного сподівання обчислюють за формулою:

$$P\{|x - m_x| < l\} = 2\Phi\left(\frac{l}{\sigma_x}\right) \cdot$$

Закон рівномірної щільності. Безперервна випадкова величина X має рівномірний розподіл на ділянці від α до β, якщо її щільність розподілу на цій ділянці постійна:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & при \quad x \in (a,b); \\ 0 & при \quad x \notin (a,b) \end{cases}.$$

Математичне сподівання: $m_x = \dfrac{b+a}{2}$; дисперсія: $D_x = \dfrac{(b-a)^2}{12}$; середнє квадратичне відхилення: $\sigma_x = \sqrt{D_x} = \dfrac{b-a}{2\sqrt{3}}$.

Імовірність влучення значень випадкової величини на інтервал (α, β):

$$P\{\alpha < X < \beta\} = \frac{\beta - \alpha}{b - a}.$$

Функція рівномірного розподілу:

$$F(x) = \frac{x-\alpha}{b-a}.$$

Задача 3.1

У відділ верхнього одягу універмагу один за одним входять три відвідувачі. За оцінками менеджера імовірність того, що відвідувач, який ввійшов, зробить покупку, дорівнює 0,3. Визначити імовірність того, що: а) жоден з відвідувачів нічого не купить; б) тільки один відвідувач зробить покупку; в) два відвідувачі зроблять покупку; г) всі троє куплять що-небудь у відділі. Побудувати ряд розподілу та визначити математичне сподівання і дисперсію випадкової величини X (число відвідувачів, що зробили покупку).

Розв'язання

Появу трьох відвідувачів у відділі універмагу можна розглядати як проведення трьох дослідів. Досліди однакові, тому що імовірність появи події A={здійснення покупки одним відвідувачем} однакова для всіх трьох і дорівнює p=0,3. Відповідно імовірність не покупки для кожного відвідувача q=0,7. Наслідки дослідів незалежні, тому що рішення про покупку для кожного з відвідувачів не залежить від рішень інших відвідувачів відділу.
Для визначення імовірностей біноміального розподілу випадкової величини X скористаємося формулою $P_n(m) = C_n^m p^m q^{n-m}$, зведемо їх до таблиці:

x_i	0	1	2	3
p_i	0,343	0,441	0,189	0,027

Обчислимо математичне сподівання та дисперсію за формулами моментів випадкової величини:

$$m_x = 0*0,343+1*0,441+2*0,189+3*0,027 = 0,9$$

$$D_x = 0*(0,343-0,9)^2+1*(0,441-0,9)^2+2*(0,189-0,9)^2+3*(0,027-0,9)^2 = \\ =0,63.$$

Обчислимо математичне сподівання і дисперсію за формулами $m_x = np$; $D_x = npq$; $\sigma_x = \sqrt{npq}$.:

$$m_x = n*p = 3 * 0,3 = 0,9$$

$$D_x = n*p*q = 3 * 0,3 * 0,7 = 0,63.$$

Задача 3.2

На АТС надходять виклики з інтенсивністю λ=0,8 1/хв. Визначити імовірність того, що протягом 2 хвилин а) не надійде жодного виклику; б) надійде рівно один виклик; в) надійде хоча б один виклик.

Розв'язання

Визначимо математичне сподівання числа викликів, що надходять на АТС, яке відповідає інтервалу часу t=2 хвилини:

$$a = \lambda * t = 0,8*2=1,6.$$

За формулою Пуассона імовірність подій визначиться в такий спосіб:

а) імовірність того, що протягом 2 хвилин не надійде жодного виклику

$$P(A) = P(0) = \frac{1{,}6^{0}}{0!} e^{-1{,}6} = 0{,}202;$$

б) імовірність того, що протягом 2 хвилин надійде рівно один виклик

$$P(B) = P(1) = \frac{1{,}6^{1}}{1!} e^{-1{,}6} = 1{,}6 * 0{,}202 = 0{,}323;$$

в) імовірність того, що протягом 2 хвилин надійде хоча б один виклик простіше визначити, використовуючи імовірність протилежної події:

$$P(C) = 1 - P(0) = 1 - 0{,}202 = 0{,}798.$$

Задача 3.3

Космічні частки, що потрапляють у супутник, утворюють поле із щільністю λ=1 частка/м^2. Агрегат супутника, який знаходиться у полі часток, займає площу s=10 см^2. Для виходу з ладу агрегату свідомо досить влучення в нього двох часток. При влученні однієї частки він виходить з ладу з імовірністю p=0,5. Визначити імовірність виходу з ладу агрегату.

Розв'язання

Позначимо подію, що цікавить нас, A={вихід агрегату з ладу}. Цій події відповідають дві гіпотези:

H_1 ={в агрегат потрапила одна частка},

H_2 ={в агрегат потрапило дві частки}.

Умовні імовірності події A: $P(A/H_1)=0{,}5$, $P(A/H_2)=1$.

Імовірності гіпотез визначимо за законом розподілу Пуассона, параметр якого a=1*0,001=0,001:

$$P(H_1) = P(1) = \frac{0{,}001^{1}}{1!} e^{-0{,}001} = 0{,}000999 \quad ,$$

$$P(H_2) = 1 - P(0) - P(1) = 1 - \frac{0{,}001^{0}}{0!} e^{-0{,}001} - P(1) = 1 - 0{,}999 - 0{,}000999 = 10^{-6}.$$

За формулою повної ймовірності дістанемо ймовірність події A:

$$P(A)=P(H_1)P(A/H_1)+ P(H_2)P(A/H_2) = 0{,}000999*0{,}5+10^{-6}*1=5{,}005*10^{-4}.$$

Задача 3.4

Є випадкова величина X з експонентним законом розподілу. Параметр розподілу λ=0,4. Визначити числові характеристики та функцію розподілу випадкової величини X, а також імовірність того, що вона прийме значення в інтервалі (6, 10).

Розв'язання

Числові характеристики випадкової величини X визначимо за формулами:

$$m_x = 1/\lambda=1/0{,}4 =2{,}5; \; D_x = 1/\lambda^2=1/(0{,}4)^2=6{,}25; \; \sigma_x = m_x = 2{,}5.$$

Щільність розподілу :

$$f(x)=\begin{cases}0, & при \quad x<0\\ 0{,}4e^{-0{,}4x}, & при \quad x\geq 0\end{cases}$$

Функція розподілу:

$$F(x)=\begin{cases}0, & при \quad x<0\\ 1-e^{-0{,}4x}, & при \quad x\geq 0\end{cases}$$

Визначимо імовірність того, що випадкова величина X прийме значення в інтервалі (6, 10):

$$P\{6\leq X\leq 10\}=e^{-\lambda\alpha}-e^{-\lambda\beta}=e^{-0{,}4*6}-e^{-0{,}4*10}=0{,}0907-0{,}0183=0{,}0724\,.$$

Задача 3.5

Відомі імовірнісні характеристики нормально розподіленої випадкової величини X: m =17; σ=0,6. Знайти імовірність події P(α<X<β); імовірність того, що P(|x-m| < δ), якщо α =16,8; β = 17,2; δ=0,3.

Розв'язання

Обчислимо імовірність, що X належить інтервалу (16,8; 17,2).

$$P(\alpha<X<\beta)=\Phi\left(\frac{\beta-m}{\sigma}\right)-\Phi\left(\frac{\alpha-m}{\sigma}\right)=$$

$$=\Phi\left(\frac{17{,}2-17}{0{,}6}\right)-\Phi\left(\frac{16{,}8-17}{0{,}6}\right)=\Phi\left(\frac{1}{3}\right)-\Phi\left(-\frac{1}{3}\right)=2*\Phi\left(\frac{1}{3}\right)=0{,}26.$$

Визначимо імовірність того, що X відхилиться від свого середнього значення m менше чим на δ:

$$P(|x-17|<0{,}3)=2*\Phi\left(\frac{0{,}3}{0{,}6}\right)=0{,}38\,.$$

Задача 3.6

Деталь, що виготовлена автоматом, вважається придатною, якщо відхилення X контрольованого розміру від номіналу не перевищує 10 мм. Точність виготовлення деталей характеризується σ=0,5. Вважаючи, що X розподілена нормально, визначити, скільки відсотків придатних деталей виготовляє автомат.

Розв'язання

Іншими словами, необхідно визначити імовірність того, що помилка X потрапить у симетричний відносно m_x інтервал, який дорівнює 10 мм.

$$P\left\{\left|\overset{o}{X}\right|<10\right\}=2*\Phi\left(\frac{l}{\sigma}\right)=2*\Phi\left(\frac{10}{5}\right)=0{,}95\,.$$

Автомат випускає 95% придатних деталей.

Задача 3.7

У нормально розподіленій сукупності 15% значень X менше за 12 та 40% значень X більше за 16,2. Знайти середнє значення та середнє квадратичне відхилення даного розподілу.

Розв'язання

З умови задачі витікає, що

P{X<12}=0,15 та P{X<16,2}=0,6.

Запишемо систему рівнянь:

$$\begin{cases} \Phi\left(\dfrac{12-m_x}{\sigma_x}\right)+0{,}5=0{,}15 \\ \Phi\left(\dfrac{16{,}2-m_x}{\sigma_x}\right)+0{,}5=0{,}6 \end{cases} \Rightarrow \begin{cases} \Phi\left(\dfrac{12-m_x}{\sigma_x}\right)=-0{,}35 \\ \Phi\left(\dfrac{16{,}2-m_x}{\sigma_x}\right)=0{,}1 \end{cases} \Rightarrow \begin{cases} \dfrac{12-m_x}{\sigma_x}=-1{,}04 \\ \dfrac{16{,}2-m_x}{\sigma_x}=0{,}25 \end{cases} \Rightarrow .$$

$$\Rightarrow \begin{cases} m_x = 15{,}386 \\ \sigma_x = 3{,}256 \end{cases}.$$

Параметри розподілу: m_x = 15,386, σ_x =3,256.

Задача 3.8

Коробки з шоколадом упаковує автомат. Їх середня маса 1,06 кг. Відомо, що 5% коробок мають масу меншу за 1 кг. Визначити відсоток коробок, маса яких перевищує 940 г.

Розв'язання

Визначимо відсоток коробок, маса яких менша за 940 г, тоді виявиться, що інші коробки мають масу, що перевищує 940 г:

$$P\{X<0{,}940\}=F(0{,}940)=\Phi\left(\frac{0{,}94-m_x}{\sigma_x}\right)+0{,}5\,.$$

Математичне сподівання відомо з умови задачі m_x = 1,06, а для визначення невідомого середнього квадратичного відхилення скористуємося тим, що за умови 5% коробок мають масу меншу за 1 кг.

$$P\{X<1{,}0\}=F(1{,}0)=\Phi\left(\frac{1{,}0-1{,}06}{\sigma_x}\right)+0{,}5=0{,}05\,,$$

звідки дістанемо:

$$\Phi\left(\frac{1{,}0-1{,}06}{\sigma_x}\right)=-0{,}45\,,\qquad \frac{1{,}0-1{,}06}{\sigma_x}=-1{,}655\,,\qquad \sigma_x=0{,}03625.$$

Тоді

$$P\{\,X<0{,}940\,\}=\Phi\left(\frac{0{,}94-1{,}06}{0{,}03625}\right)+0{,}5=\Phi(-3{,}31)+0{,}5=-0{,}499+0{,}5=0{,}001\,.$$

Відсоток коробок, маса яких перевищує 940 г становить 99,9%.

Задача 3.9

Довжину кімнати вимірюють рулеткою із грубими діленнями (10 см). Округлення провадять до найближчого цілого. X - помилка вимірювання. Знайти її щільність розподілу, функцію розподілу та числові характеристики.

Розв'язання

Довжина кімнати L з урахуванням помилки визначиться як L±5 см, тобто випадкова величина X змінюється в межах -5<X<+5. Оскільки крива щільності розподілу обмежує площу, що дорівнює одиниці, значення f(x) дістанемо в такий спосіб:

$$f(x) = \frac{1}{b-a} = \frac{1}{5-(-5)} = 0{,}1.$$

Математичне сподівання:

$$m_x = \frac{a+b}{2} = \frac{-5+5}{2} = 0.$$

Дисперсія:

$$D_x = \frac{(b-a)^2}{12} = \frac{(5+5)^2}{12} = 8{,}33; \qquad \sigma x = \sqrt{8{,}33} = 2{,}89.$$

Задача 3.10

Поїзда метро йдуть регулярно з інтервалом 2 хвилини. Пасажир виходить на платформу у випадковий момент часу, не пов'язаний з розкладом поїздів. Випадкова величина T - час очікування поїзда. Знайти: а) щільність розподілу і числові характеристики випадкової величини T; б) імовірність того, що чекати доведеться не більше 0,5 хвилини.

Розв'язання

Оскільки поїзди під'їжджають до станції рівномірно, закон розподілу випадкової величини T – рівномірний, тобто $f(x) = \frac{1}{b-a}$, де (b-a) – інтервал руху поїздів, причому a =0, b =2, тоді:

а) $f(x) = \frac{1}{2-0} = 0{,}5$; $m_x = \frac{b+a}{2} = \frac{2+0}{2} = 1$ хв.; $D_x = \frac{(b-a)^2}{12} = \frac{(2-0)^2}{12} = \frac{1}{3}$.

б) імовірність того, що чекати доведеться не більше 0,5 хвилини визначимо так:

$$P\{T < 0{,}5\} = F(x) = \frac{0{,}5-0}{2-0} = 0{,}25.$$

Практичне заняття 4
СИСТЕМА ВИПАДКОВИХ ВЕЛИЧИН.
ЗАКОНИ РОЗПОДІЛУ ТА ЧИСЛОВІ ХАРАКТЕРИСТИКИ СИСТЕМИ

Мета — сформувати вміння користуватися законами розподілу системи двох випадкових величин для визначення числових характеристик системи, визначення їх залежності, та вміння обчислювати числові характеристики функцій випадкових величин.

Для характеристики системи двох випадкових величин використовують закони розподілу системи та числові характеристики системи. Найпростішим є закон розподілу системи двох дискретних випадкових величин, що уявляє собою кореляційну таблицю, у якій перший рядок містить всі значення випадкової величини X, а перший стовпець - всі значення випадкової величини Y. В ij-ї клітині таблиці записують імовірність події $\{X=x_i,\ Y=y_j\}$. Сума всіх імовірностей у кореляційній таблиці дорівнює одиниці

$$\sum_{i=1}^{n}\ \sum_{j=1}^{m} p\{x_i, y_j\} = 1 .$$

Функція розподілу системи двох випадкових величин (X,Y) дорівнює імовірності того, що випадкова величина X прийме значення, менше за x та випадкова величина Y прийме значення, менше за y: $F(x,y) = P\{X < x,\ Y < y\}$.

Властивості функції розподілу системи:

1. Функція розподілу - неубутна функція, тобто $F(x_2,y) \geq F(x_1,y)$, якщо $x_2 > x_1$; $F(x,y_2) \geq F(x,y_1)$, якщо $y_2 > y_1$.

2. Функція розподілу дорівнює нулю, якщо хоча б один з аргументів обертається на мінус нескінченність: $F(x, -\infty) = F(-\infty, y) = F(-\infty, -\infty) = 0$.

3. Функція розподілу, якщо хоча б один з аргументів обертається на плюс нескінченність, дорівнює функції розподілу компоненту системи, що залишився: $F(x, +\infty) = F(x)$; $F(+\infty, y) = F(y)$; $F(+\infty, +\infty) = 1$.

Щільність розподілу імовірностей системи двох випадкових величин (X,Y) $f(x,y)$ являє собою другу змішану похідну від $F(x,y)$:

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} .$$

Тут припускається, що $F(x,y)$ є безперервною та двічі диференційованою. Властивості щільності розподілу імовірностей:

1. Щільність розподілу невід'ємна, тобто $f(x,y) \geq 0$;

2. Функція розподілу системи двох випадкових величин (X,Y) дорівнює подвійному інтегралу від щільності розподілу системи:

$$F(x,y) = \iint\limits_{-\infty-\infty}^{xy} f(x,y)dxdy ,$$

де $f(x,y)dxdy$ - елемент імовірності, що є імовірністю влучення системи в елементарний прямокутник dxdy та дорівнює об'єму паралелепіпеда f(x)dxdy.

3. Подвійний інтеграл від щільності розподілу у нескінченних межах дорівнює одиниці:

$$\iint\limits_{-\infty}^{+\infty} f(x,y)dxdy = 1 .$$

Математичні сподівання випадкових компонентів системи X та Y: $\alpha_{1,0} = M[X^1Y^0] = M[X]$; $\alpha_{0,1} = M[X^0Y^1] = M[Y]$.

Дисперсії випадкових компонентів системи X та Y: $\mu_{2,0} = M[\overset{0}{X}{}^{2}\,\overset{0}{Y}{}^{0}] = M[\overset{0}{X}{}^{2}] = D_x$; $\mu_{0,2} = M[\overset{0}{X}{}^{0}\,\overset{0}{Y}{}^{2}] = M[\overset{0}{Y}{}^{2}] = D_y$.

Кореляційний момент для дискретної випадкової величини визначається за формулою: $K_{x,y} = \sum_{i=1}^{n} \sum_{j=1}^{m} \overset{0}{x_i} \overset{0}{y_j} p_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{m} (x_i - m_x)(y_j - m_y) p_{ij}$, де $p_{ij} = P\{X=x_i|Y=y_j\}$ – умовна імовірність, тобто імовірність того, що X прийме значення x_i за умови, що Y прийме значення y_j. Для безперервної випадкової величини:

$$K_{xy} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x - m_x)(y - m_y) f(x,y) dxdy .$$

Кореляційний момент є характеристикою зв'язку між величинами X та Y, і у випадку незалежних X та Y він дорівнює нулю.

Для визначення тісноти зв'язку між X та Y використовують *коефіцієнт кореляції* r_{xy}, який визначається за формулою:

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} .$$

Функції випадкових величин.

1. Математичне сподівання суми двох залежних або незалежних випадкових величин X та Y дорівнює сумі їх математичних сподівань:

$$M[X+Y] = \Sigma(x_i+y_i)p_i = \Sigma x_i p_i + \Sigma y_i p_i = M[X] + M[Y].$$

Для n доданків: $M[\Sigma X_i] = \Sigma M[X_i]$.

Математичне сподівання лінійної функції кількох випадкових величин $Y = \sum_{i=1}^{n} a_i X_i + b$ дорівнює тієї самій лінійній функції від їх математичних сподівань: $M[Y] = M\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} a_i M[X_i] + b$, де a_i і b_i – невипадкові коефіцієнти.

2. Математичне сподівання добутку двох випадкових величин X та Y дорівнює добутку їх математичних сподівань плюс кореляційний момент:

$$M[XY] = M[X] * M[Y] + K_{xy}.$$

Якщо випадкові величини X та Y некорельовані, математичне сподівання добутку дорівнює добутку їх математичних сподівань: $M[XY] = M[X] * M[Y]$, або для n незалежних співмножників: $M\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} M[X_i]$.

3. Дисперсія суми двох випадкових величин X та Y дорівнює сумі їх дисперсій плюс подвоєний кореляційний момент $D[X+Y] = D[X] + D[Y] + 2K_{xy}$.

Дисперсія суми кількох випадкових величин виражається формулою $D\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} D[X_i] + 2\sum_{i<j} K_{x_i x_j}$, де $K_{x_i x_j}$ - кореляційний момент випадкових величин X_i та X_j.

Якщо X_i - некорельовані випадкові величини, дисперсія їх суми дорівнює сумі їх дисперсій, тоді для суми n незалежних випадкових величин дисперсія визначиться в такий спосіб: $D[\Sigma X_i] = \Sigma D[X_i]$, звідки середнє квадратичне відхилення суми $\sigma_\Sigma = \sqrt{\Sigma \sigma_i^2}$.

Дисперсія лінійної функції кількох випадкових величин $Y = \sum_{i=1}^{n} a_i X_i + b$ виражається формулою $D[Y] = D\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} a_i^2 D[X_i] + 2\sum_{i<j} a_i a_j K_{x_i x_j}$.

Якщо X_i - некорельовані випадкові величини, то $D[Y] = D\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} a_i^2 D[X_i]$.

4. Дисперсію добутку двох незалежних випадкових величин X та Y обчислюють за формулою $D[XY] = D[X]*D[Y] + M[X]^2 D[Y] + M[Y]^2 D[X]$.

Задача 4.1

Матриця розподілу випадкового вектора (X, Y) має вигляд:

y_i / x_i	0	2	5
1	0,1	0	0,2
2	0	0,3	0
4	0,1	0,3	0

Знайти числові характеристики.

Розв'язання

Для визначення числових характеристик випадкової величини X побудуємо її ряд розподілу.

x_i	1	2	4
p_i	0,3	0,3	0,4

де P{X=1}=0,1+0+0,2=0,3; P{X=2}=0+0,3+0=0,3; P{X=4}=0,1+0,3+0=0,4.

Тоді $m_x = \Sigma x_i p_i = 1*0,3+2*0,3+4*0,4=2,5$;

$D_x = \Sigma(x_i - m_x)^2 p_i =(1-2,5)^2*0,3+(2-2,5)^2*0,3+(4-2,5)^2*0,4=1,65$;

$\sigma_x = \sqrt{D_x} = \sqrt{1,65} = 1,285$.

Аналогічно побудуємо ряд розподілу випадкової величини Y.

y_i	0	2	5
p_i	0,2	0,6	0,2

та визначимо її числові характеристики:

$m_y = \Sigma y_i p_i = 0*0,2+2*0,6+5*0,2=2,2$;

$D_y = \Sigma(y_i - m_y)^2 p_i =(0-2,2)^2*0,3+(2-2,2)^2*0,6+(5-2,2)^2*0,2=2,54$;

$\sigma_y = \sqrt{D_y} = \sqrt{2,54} = 1,6$.

Визначимо кореляційний момент системи випадкових величин X, Y:

$$K_{xy} = \sum_{i=1}^{3} \sum_{j=1}^{3} (x_i - m_x)(y_j - m_y) p_{ij} = (1-2,5)[(0-2,2)*0,1 + (2-2,2)*0 + (5-2,2)*0,2] +$$

$$+(2-2,5)[(0-2,2)*0+(2-2,2)*0,3+(5-2,2)*0]+$$
$$+(4-2,5)[(0-2,2)*0,1+(2-2,2)*0,3+(5-2,2)*0]=-0,9.$$

Визначимо коефіцієнт кореляції:

$$r_{xy}=\frac{K_{xy}}{\sigma_x\sigma_y}=\frac{-0,9}{1,285*1,6}=-0,438.$$

Задача 4.2

Двічі кидають гральну кістку. Випадкові величини X - число появ шістки, Y - число появ парної цифри. Описати закони розподілу випадкових величин X та Y; описати закон розподілу випадкового вектора (X, Y); встановити, чи є залежними X та Y.

Розв'язання

Знайдемо закон розподілу випадкового вектора (X, Y), склавши матрицю розподілу, оскільки X та Y - дискретні.

y_i / x_i	0	1	2	$P\{X=x_i\}$
0	1/4	1/3	1/9	25/36
1	0	1/6	1/9	10/36
2	0	0	1/36	1/36
$P\{Y=y_i\}$	1/4	1/2	1/4	1

Для визначення умовних імовірностей $P_{ij}=P\{X=x_i/Y=y_j\}$ скористуємося теоремами додавання та множення імовірностей.

Дістанемо ряд розподілу випадкової величини X, для чого підсумовуємо умовні імовірності у кореляційній таблиці за рядками, та Y, для чого підсумовуємо умовні імовірності за стовпцями.

Встановимо, чи залежні X та Y. Відомо, що у випадку незалежних подій умовна імовірність події дорівнює її безумовній імовірності, але рівність $P\{y_j/x_i\}=P\{y_j\}$ не дотримується. Наприклад,

$$P_{22}\{y_j=1,\ x_i=1\}=P\{x_i=1\}*P\{y_j=1/x_i=1\}=1/6.$$

Отже, X та Y залежні.

Задача 4.3

Для умови попередньої задачі визначити числові характеристики випадкового вектора (X, Y).

Розв'язання

Визначимо математичні сподівання X та Y:

$$m_x=\Sigma x_i p_i=0*\frac{25}{36}+1*\frac{5}{18}+2*\frac{1}{36}=\frac{1}{3};$$

$$m_y=\Sigma y_i p_i=0*\frac{1}{4}+1*\frac{1}{2}+2*\frac{1}{4}=1.$$

Визначимо дисперсії X та Y:

$$D_x = \Sigma(x_i - m_x)^2 p_i = \frac{5}{18} \;;$$

$$D_y = \Sigma(y_i - m_y)^2 p_i = \frac{1}{2} \;.$$

Визначимо кореляційний момент системи випадкових величин X, Y:

$$K_{xy} = \sum_{i=1}^{3} \sum_{j=1}^{3} (x_i - m_x)(y_j - m_y)\, p_{ij} = (0 - \frac{1}{3})\left[(0-1)*\frac{1}{4} + (1-1)*\frac{1}{3} + (2-1)*\frac{1}{9}\right] +$$

$$+ (1 - \frac{1}{3})\left[0 + (1-1)*\frac{1}{6} + (2-1)*\frac{1}{9}\right] +$$

$$+ (2 - \frac{1}{3})\left[(0-1)*0 + (1-1)*0 + (2-1)*\frac{1}{36}\right] = \frac{1}{6}.$$

Визначимо коефіцієнт кореляції:

$$r_{xy} = \frac{K_{xy}}{\sqrt{D_x D_y}} = \frac{\frac{1}{6}}{\sqrt{\frac{5}{18} * \frac{1}{6}}} = \frac{1}{\sqrt{5}} .$$

Задача 4.4

На вимірювальний прибор надходить випадковий вектор (X, Y) з наступними характеристиками: m_x=-1; m_y=1; σ_x=2; σ_y=3; r_{xy}=0,5. На виході прибору вимірюють величину $Z=(X-Y)^2$. Визначити математичне сподівання випадкової величини Z.

Розв'язання

Оскільки коефіцієнт кореляції не дорівнює нулю, випадкові величини X та Y корельовані. Перетворимо вираз Z:

$$M[Z] = M[(X\text{-}Y)^2] = M[X^2\text{-}2XY+Y^2]=$$

математичне сподівання суми випадкових величин дорівнює сумі математичних сподівань:

$$= M[X^2]\text{-}\ M[2XY]+ M[Y^2]=$$

другий початковий момент запишемо як суму дисперсії та квадрата математичного сподівання і врахуємо, що математичне сподівання добутку корельованих випадкових величин дорівнює добутку їх математичних сподівань плюс кореляційний момент:

$$= D_x+M^2[X] - 2(M[X]*M[Y]+K_{xy})+ D_y+M^2[Y] =$$

$$= 2^2+(-1)^2 - 2*(-1*1+2*3*0,5) + 3^2+1^2 = 11.$$

Задача 4.5

Визначити математичне сподівання та дисперсію випадкової величини Z, якщо Z=3+4(X-Y). Числові характеристики: m_x=-2; m_y=4; D_x=4; D_y=9; r_{xy}=-0,5.

Розв'язання

Оскільки коефіцієнт кореляції не дорівнює нулю, випадкові величини X та Y корельовані. Визначимо математичне сподівання Z:

$$M[Z] = M[3+4(X-Y)] = M[3]+4M[X+Y]=$$
$$= 3+4m_x+4m_y=3+4*(-2)+4*4=11.$$

Визначимо дисперсію Z:

$$D[Z]= D[3+4(X-Y)] = D[3]+16D[X+Y]=$$
$$= 0+16(D[X]+D[Y]+2K_{xy}) =$$
$$=16*[4+9+2*(-0,5)(4*9] = 112,$$

де $K_{xy}=r_{xy}*\sqrt{D_x D_y}$.

Середнє квадратичне відхилення $\sigma_z = \sqrt{D_z} = \sqrt{112} = 10,6$.

Задача 4.6

Випадкові величини X та Y незалежні і мають наступні характеристики: $m_x=1$; $m_y=2$; $\sigma_x=1$; $\sigma_y=2$. Обчислити математичне сподівання випадкових величин:

а) $U=X^2+2Y^2-XY-4X+Y+4$;

б) $V=(X+Y-1)^2$.

Розв'язання

а) Визначимо математичне сподівання випадкової величини U:

$$M[U]=M[X^2]+2M[Y^2]-M[XY]-4M[X]+M[Y]+M[4] =$$
$$= D[X]+M^2[X]+2(D[Y]+M_2[Y])-M[X]*M[Y]-4M[X]+M[Y]+4=$$
$$1+1+2*(4+4)-1*2-4*1+2+4 = 18;$$

б) Визначимо математичне сподівання випадкової величини V:

$$M[V] = M[(X+Y-1)^2]= M[X^2+2Y-2+Y^2-2Y+1] =$$
$$= M[X^2]+2M[Y]-M[2]+M[Y^2]-2M[Y]+M[1] =$$
$$= D[X]+M^2[X]+2M[Y]-2+D[Y]+M^2[Y]-2M[Y]+1=$$
$$=1+1+4-2+4+4-4+1 = 9.$$

Задача 4.7

Є випадкова величина X з математичним сподіванням m_x та дисперсією D_x. Знайти математичне сподівання та дисперсію наступних випадкових величин:

а) $Y=-X$;

б) $Z=X+2Y-1$;

в) $U=3X-Y+2Z-3$.

Розв'язання

а) Визначимо математичне сподівання випадкової величини Y:

а) $M[Y]=M[-X]= -m_x$;

визначимо дисперсію випадкової величини Y

$$D[Y] = D[-X] = (-1)^2D_x = D_x;$$

б) Визначимо математичне сподівання випадкової величини Z:

$$M[Z]=M[X+2Y-1] = m_x -2m_x -1 = - m_x -1;$$

визначимо дисперсію випадкової величини Z з урахуванням того, що математичне сподівання добутку корельованих випадкових величин дорівнює добутку їх математичних сподівань плюс кореляційний момент:

$$D[Z]=D[X+2Y-1] = D_x + 4D_y + 2Kxy =$$

кореляційний момент запишемо як різницю другого початкового моменту і добутку математичних сподівань, дістанемо

$$K_{xy} = M[X2Y] - M[X] * M[2Y]= M[X(-2X)] - M[X] * M[2Y] =$$
$$= -2M[X^2] + 2M^2[X] = -2(M[X^2] - M^2[X]) = -2D_x.$$

тоді дисперсія Z дорівнюватиме

$$D[Z] = D_x + 4D_x – 4D[X] = D[X].$$

в) Визначимо математичне сподівання випадкової величини Z:

$$M[U]= M[3X-Y+2Z-3] = 3m_x – m_y +2m_z -3 =$$
$$=3m_x + m_x +2(- m_x -1) -3 = 2m_x -5;$$

визначимо дисперсію випадкової величини U

$$D[U]= D[3X-Y+2Z-3] = D[3X] + D[-Y] + D[2Z] + 2(K_{xy}+K_{yz} + K_{xz});$$

визначимо кореляційні моменти

$$K_{xy} = M[3X(-Y)] - M[3X] * M[-Y]= 3M[X^2] - 3M[X] * M[X]=$$
$$= 3M[X^2] - 3M^2[X] = 3D_x;$$
$$K_{xz} = M[3X2Z] - M[3X] * M[2Z]= M[3X2(X+2Y-1)] - M[3X] * M[2X+4Y-2]=$$
$$= M[6X^2 -12X^2-6X)] - 3M[X] *(2M[X]+4M[-X]-2) =$$
$$=M[6X^2] - 12M[X^2]- 6M[X] - 6M^2[X] +12M^2[X] + 6 M[X] =$$
$$=-6 M[X^2] + 6 M^2[X] = -6 (M[X^2] - M^2[X]) = -6D_x.$$
$$K_{yz} = M[-Y2Z] - M[-Y] * M[2Z]= M[2X (X+2Y-1)] - M[X] * M[2(X+2Y-1)]=$$
$$= M[2X^2 -4X^2-2X] - M[X] * (2M[X]-4M[X]-2)=$$
$$=2M[X^2] -4M[X^2]-2M[X] - 2M^2[X] +4M^2[X]+2M[X]=$$
$$=-2M[X^2] + 2M^2[X]=-2(M[X^2] - M^2[X]) = -2D_x.$$

Визначимо тепер дисперсію випадкової величини U:

$$D[U]= D[3X-Y+2Z-3] = 9D[X] + D[X] + 4D[X] + 2(3D_x-6D_x-2D_x)=$$
$$=9D_x + D_x + 4D_x + 6D_x-12D_x-4D_x= 4 D_x.$$

Практичне заняття 5

ОБРОБКА СТАТИСТИЧНИХ ДАНИХ. ПОБУДОВА ЗАКОНУ РОЗПОДІЛУ ВИПАДКОВОЇ ВЕЛИЧИНИ ЗА СТАТИСТИЧНИМИ ДАНИМИ

Мета — сформувати вміння побудувати закон розподілу випадкової величини на підставі статистичних даних та визначати статистичні оцінки параметрів розподілу.

Варіаційним рядом називають таблицю, в одному рядку якої розташовуються варіанти $x_1, x_2, \ldots, x_n$ у зростаючому або убутному порядку, а в другому — відповідні їм частоти $m_1, m_2, \ldots, m_n$...

x_i	x_1	x_2	…	x_n
m_i	m_1	m_2	…	m_n

де $\sum_{i=1}^{n} m_i = n$.

Дискретною називають варіацію, при якій окремі значення ознаки (варіанти) відрізняються одна від одної на певну скінчену величину.

Безперервною називають варіацію, при якій значення ознаки можуть відрізнятися одне від одного на як завгодно малу величину.

Замість абсолютних частот m_i зазвичай використовують відносні – p^*_i. Для одержання відносних частот необхідно відповідну частоту розділити на суму всіх частот: $p_1^* = \dfrac{m_1}{\sum_{i=1}^{n} m_i}$, $p_2^* = \dfrac{m_2}{\sum_{i=1}^{n} m_i}$, …, $p_n^* = \dfrac{m_n}{\sum_{i=1}^{n} m_i}$. Сума всіх відносних частот дорівнює одиниці: $\sum_{i=1}^{n} p_i^* = 1$.

Полігон розподілу будують у прямокутній системі координат. Величину ознаки відкладають на осі абсцис, частоти або відносні частоти - на осі ординат.

Кумулятивна крива (кумулята) утворюється при зображенні варіаційного ряду з накопиченими відносними частотами у прямокутній системі координат. Накопичена частота певної варіанти утворюється підсумовуванням всіх частот варіант, які передують даній, із частотою цієї варіанти. При $\Delta x_i \to 0$ кумулята прагне до безперервної кривої, яка є статистичною функцію розподілу досліджуваної ознаки.

Гістограму розподілу будують аналогічно полігону у прямокутній системі координат. При побудові гістограми на осі абсцис вибирають відрізки, що відповідають інтервалам, на яких будують прямокутники із площею, пропорційною відносним частотам інтервалів p_i^*. З умови побудови гістограми випливає, що вся її площа дорівнює 1. При $\Delta x_i \to 0$ гістограма прагне до безперервної залежності, що є статистичною щільністю розподілу випадкової величини.

Однією з задач математичної статистики є отримання оцінок числових характеристик досліджуваної ознаки. Іншою задачею є визначення закону розподілу досліджуваної ознаки за статистичним даними. При цьому виникає задача із згладжування статистичних рядів за допомогою аналітичних виразів.

Щоб знайти статистичну функцію розподілу $F^*(x)$, підраховують число варіант, в яких ознака X прийняла значення менші за x, тобто $X<x$. Якщо число таких варіант $m(x)$, а обсяг сукупності дорівнює n, то $F^*(x) = \dfrac{m(x)}{n}$.

Якщо результати спостереження зведені у групований статистичний ряд, на його підставі будують гістограму, з урахуванням того, що частота появ ознаки на і-му інтервалі $p^* = \frac{m_i}{n}$, де m_i – кількість появ х на і-му інтервалі.

Для оформлення статистичного ряду у вигляді гістограми на осі абсцис відкладають інтервали ряду, а потім на кожному інтервалі будують прямокутник, площа якого дорівнює p^*_i.

Задача вирівнювання статистичних рядів полягає у знаходженні теоретичної кривої розподілу. Якщо клас функцій, що описують розподіл, відомий, то задача зводиться до раціонального вибору параметрів розподілу. Одним з методів розв'язання цієї задачі є метод моментів, відповідно до якого числові параметри розподілу вибирають так, щоб найважливіші числові характеристики дорівнювали їх статистичним оцінкам. Зокрема, математичне сподівання та дисперсія приймаються рівними їх статистичним оцінкам - вибірковій середній та вибірковій дисперсії: $m_x = \tilde{X}$, $D_x = \sigma^2_{виб}$. Інший метод визначення статистичних оцінок параметрів розподілу є метод максимальної правдоподібності.

Числові характеристики варіаційного ряду. Однією з найважливіших характеристик варіаційного ряду є його *середнє значення*, яку визначають як відношення суми добутків варіантів на відповідні частоти до суми всіх частот:

$$\overline{x} = \frac{\sum_{i=1}^{k} x_i m_i}{\sum_{i=1}^{k} m_i}, \quad \overline{x} = \sum_{i=1}^{k} x_i p_i^*, \text{ або } \overline{x} = \frac{\sum_{i=1}^{n} x_i}{n},$$

де m_i – частоти варіаційного ряду; p_i^* - відносні частоти; k – число груп.

Дисперсія варіаційного ряду – це середня арифметична квадрата відхилення значень ознак ряду від їх середньої арифметичної. Дисперсія обчислюється за формулами:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{X})^2}{n} \qquad \sigma^2 = \frac{\sum_{i=1}^{k}(x_i - \overline{X})^2 m_i}{\sum_{i=1}^{k} m_i}$$

Стандартне відхилення варіаційного ряду визначають як арифметичне значення квадратного кореня з дисперсії.

$$\sigma = \sqrt{\sigma^2} .$$

Задача 5.1

За результатами розпродажу товарів 26 продавцями побудований варіаційний ряд:

Число продажів (x_i)	9	12	13	14	15	16	17	19	21	23	27
Число продав-ців(m_i)	1	2	3	6	5	3	2	1	1	1	1

Знайти статистичну функцію розподілу F*(x).

Розв'язання

Очевидно, що шукана статистична функція розподілу для всіх x<9 дорівнює нулю (тому що число продажів не приймало значення менше за дев'ять)

F*(9)=0.

Визначимо частоту числа продажів x<12, до них належить число продажів 9, причому m(9)=1, тоді

$$F^*(12) = \frac{m(9)}{n} = \frac{1}{26} \approx 0{,}04\,.$$

Визначимо частоту числа продажів x<13, до них належать число продажів 9 та 12, причому m(9)=1, а m(12)=2, тоді

$$F^*(13) = \frac{m(9)}{n} + \frac{m(12)}{n} = \frac{1}{26} + \frac{2}{26} \approx 0{,}04 + 0{,}08 = 0{,}12\,.$$

Аналогічно отримаємо інші значення статистичної функції розподілу та подамо їх у табличному вигляді.

x_i	9	12	13	14	15	16	17	19	21	23	27	x>27
F(x_i)	0	0,04	0,12	0,23	0,46	0,65	0,77	0,85	0,88	0,92	0,96	1

Графік статистичної функції розподілу представлений на рис. 5.1.

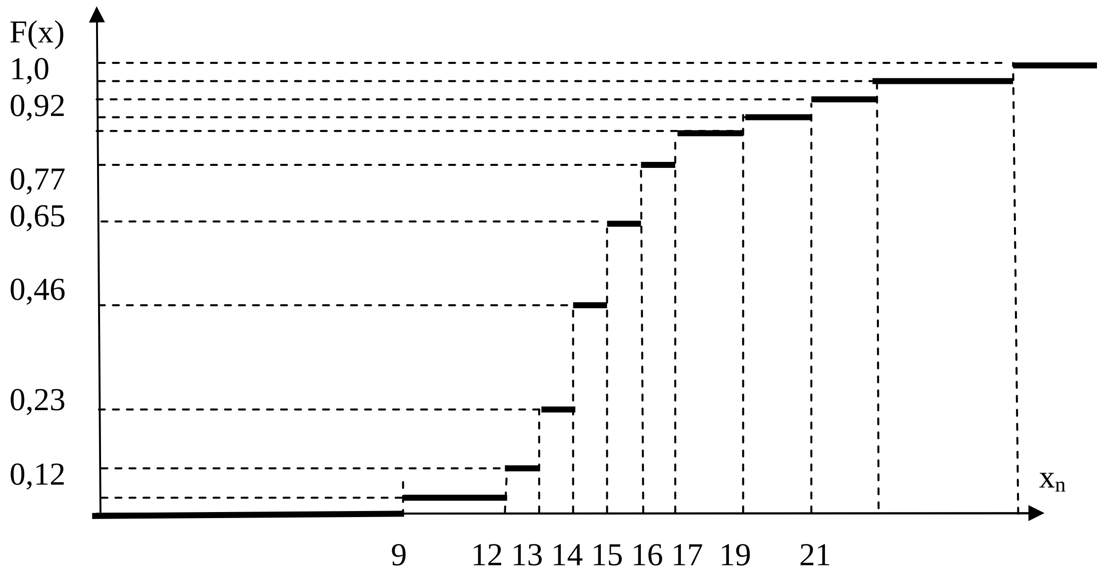


Рис. 5.1 - Статистична функція розподілу.

Задача 5.2

З метою дослідження точності прибору зробили 500 вимірювань помилки. Результати вимірювань звели у групований варіаційний ряд:

групи	-4 ...-3	-3 ...-2	-2 ...-1	-1 ...0	0 ...1	1 ...2	2 ...3	3 ...4
m_i	6	25	72	133	120	88	46	10

Побудувати гістограму розподілу та визначити параметри розподілу.

Розв'язання

Визначимо частоти для кожного розряду групованого варіаційного ряду, користуючись формулою

$$p_i^* = \frac{m_i}{n},$$

де m_i – число значень помилки X, що потрапили в i-у групу; n – число зроблених вимірів, n = 500.

групи	-4 …-3	-3 …-2	-2 …-1	-1 …0	0 …1	1 …2	2 …3	3 …4
m_i	6	25	72	133	120	88	46	10
p_i*	0,012	0,05	0,144	0,266	0,24	0,176	0,092	0,02

Для побудови гістограми визначимо значення щільності частот для кожної групи значень за формулою

$$f_i^* = \frac{p_i{}^*}{l},$$

де l - довжина групи, l= 1.

групи	-4 …-3	-3 …-2	-2 …-1	-1 …0	0 …1	1 …2	2 …3	3 …4
m_i	6	25	72	133	120	88	46	10
p_i*	0,012	0,05	0,144	0,266	0,24	0,176	0,092	0,02
f_i*	0,012	0,05	0,144	0,266	0,24	0,176	0,092	0,02

Побудуємо графік гістограми (рис. 5.2).

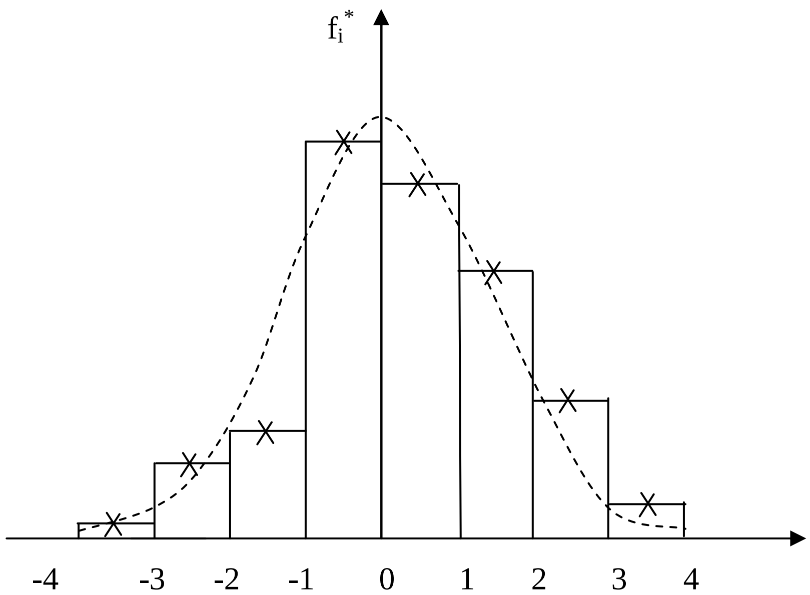


Рис. 5.2 - Графік гістограми

З вигляду гістограми можна припустити, що її можна згладити за допомогою нормального закону (припущення підтверджується тим, що досліджувана випадкова величина є помилкою вимірювання, а отже розподілена нормально), щільність розподілу якого визначають за виразом

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}.$$

Параметри m та σ, що входять у вираз щільності f(x), виберемо так, щоб щонайкраще погодити аналітичний вираз зі статистичним розподілом. Визначимо вибіркове середнє $\tilde{X}$ та вибіркове середнє квадратичне відхилення $\sigma_{выб}$ за даними групованого варіаційного ряду. Як значення x_i виберемо середину i-ї групи і цьому значенню поставимо у відповідність як імовірність його частоту p_i^*, отримаємо

$$\tilde{X}=\sum_{i=1}^{8}x_i p_i^*=-3{,}5*0{,}012-2{,}5*0{,}05-1{,}5*0{,}144-0{,}5*0{,}266+\\+0{,}5*0{,}24+1{,}5*0{,}176+2{,}5*0{,}092+3{,}5*0{,}02=0{,}168;$$

$$\tilde{\alpha}_2=\sum_{i=1}^{8}x_i^2 p_i^*=(-3{,}5)^2*0{,}012+(-2{,}5)^2*0{,}05+(-1{,}5)^2*0{,}144+(-0{,}5)^2*0{,}266+\\+(0{,}5)^2*0{,}24+(1{,}5)^2*0{,}176+(2{,}5)^2*0{,}092+(3{,}5)^2*0{,}02=2{,}126;$$

$$\sigma_{выб}^2=\tilde{\alpha}_2-\tilde{X}^2=2{,}126-(0{,}168)^2=2{,}098;$$

$$\sigma_{выб}=\sqrt{\sigma_{выб}^2}=\sqrt{2{,}098}=1{,}448.$$

Дістали розподіл $f^*(x)=\frac{1}{1{,}448\sqrt{2\pi}}\exp\left\{-\frac{(x-0{,}168)^2}{2*2{,}126}\right\}$, користуючись яким, підрахуємо значення f*(x) на межах груп:

$f_i(-4)=0{,}0045$; $f_i(1)=0{,}2343$;
$f_i(-3)=0{,}0256$; $f_i(2)=0{,}1244$;
$f_i(-2)=0{,}0895$; $f_i(3)=0{,}0435$;
$f_i(-1)=0{,}1986$; $f_i(4)=0{,}0087$;
$f_i(0)=0{,}274$.

Відкладемо на графіку отримані точки та проведемо плавну криву (рис. 5.2).

Практичне заняття 6
ВЛАСТИВОСТІ ВИБІРКОВИХ ЧИСЛОВИХ ХАРАКТЕРИСТИК

Мета — сформувати вміння оцінювати властивості статистичних оцінок числових параметрів розподілу та визначати можливі похибки та їх імовірності, а також щонайменшу кількість дослідів для їх запобігання.

Оцінка параметра a^* є *спроможною*, якщо при $n \to \infty$ вона збігається за імовірністю до оцінюваного параметра a: $\lim_{n\to\infty} a^* = a$.

Оцінка параметра a^* є *незміщеною* (тобто не містить систематичної помилки), якщо її математичне сподівання дорівнює оцінюваному параметру a: $M[a^*] = a$.

Оцінка параметра a^* є *ефективною*, якщо при заданому обсязі вибірки вона має найменшу дисперсію. Ступінь ефективності оцінюють відношенням дисперсій: $F = \frac{\sigma_1^2}{\sigma_2^2}$, тобто якщо F>1, то σ^2_2 ефективніша, і навпаки.

Математичне сподівання вибіркової середньої не залежить від числа дослідів n і дорівнює генеральній середній: $M[\tilde{X}] = \overline{X}$.

Дисперсія вибіркової середньої: $\sigma^2[\tilde{X}] = \frac{\sigma^2_{ген}}{n}$.

Середнє квадратичне відхилення вибіркової середньої: $\sigma[\tilde{X}] = \frac{\sigma_{ген}}{\sqrt{n}}$.

Отже, вибіркова середня є незміщеною оцінкою генеральної середньої.

Незміщену оцінку дисперсії позначають S^2 та визначають за формулою:

$$S^2 = \frac{n}{n-1}\sigma^2_{выб} = \frac{\sum\left(x_i - \tilde{X}\right)^2}{n-1}.$$

Оцінку кореляційного моменту визначають за формулою:

$$K_{хувув} = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{n-1}.$$

а оцінку коефіцієнта кореляції - за формулою

$$r_{хувув} = \frac{K_{хувув}}{\sigma_{выб_x}\sigma_{выб_y}}.$$

Приведені оцінки є спроможними та незміщеними.

Довірчий інтервал і довірча імовірність. Якщо точкову оцінку параметра визначено на підставі вибірки малого обсягу, вона може значно відрізнятися від оцінюваного параметра. Для визначення помилки від заміни генерального параметра його оцінкою використовують поняття довірчого інтервалу та довірчої імовірності.

При визначенні дійсного значення шуканого параметра m_x величину погрішності характеризує довірчий інтервал L, а ступінь впевненості, що похибка не перевищить L, характеризує довірча імовірність β.

Якщо для певного параметру розподілу, наприклад, математичного сподівання m_x отримано спроможну та незміщену оцінку a^*, потрібно знати, до яких помилок може призвести заміна параметра m_x його точковою оцінкою, і з яким ступенем впевненості можна очікувати, що ці помилки не вийдуть за певні межі. Призначають досить велику імовірність β (0,95; 0,99) таку, що подію A=$\{|a^*-m_x|<l\}$, яка характеризується цією імовірністю, можна вважати практично вірогідною, потім знаходять таке значення l, для якого справедлива рівність

$$P(A) = P\{(a- l)< m_x <(a+ l)\} = \beta.$$

Тобто з імовірністю β невідоме значення параметра m_x буде перебувати в інтервалі L=[a^*- l, a^* + l]. Більші за абсолютним значенням помилки, зустрічатимуться з імовірністю α= 1 - β. Границі інтервалу називають довірчими границями: $a_1 = a^*$-l; $a_2 = a^*$+l.

Утруднення полягає в тому, що закон розподілу a^* залежить від закону розподілу досліджуваної ознаки X і, отже, від його невідомих параметрів (зокрема і від самого параметра a). Щоб обійти це утруднення, застосовують наступний грубо наближений прийом: замінюють у виразі для l невідомі параметри їх точковими значеннями. При 20-30 дослідах цей прийом зазвичай дає задовільні за точністю результати.

Задача 6.1

Для визначення точності вимірювального прибору було зроблено п'ять незалежних вимірювань, результати яких зведені у таблицю:

Номер вимірювання	1	2	3	4	5
x_i	2781	2836	2807	2763	2858

Визначити незміщену оцінку дисперсії помилок вимірювального приладу, якщо дійсне значення вимірюваної величини:

а) відомо і дорівнює 2800; б) невідомо.

Розв'язання

а) якщо значення вимірюваної величини відомо, то генеральна середня $\overline{X}$ = 2800, незміщену оцінку дисперсії у цьому випадку можна визначити за формулою

$$\sigma^2_{выб} = \frac{\sum(x_i-\overline{x})^2}{n} = \frac{(2781-2800)^2+(2836-2800)^2+(2807-2800)^2+(2763-2800)^2+(2858-2800)^2}{5} =$$

$$=1287{,}8\,.$$

б) якщо значення вимірюваної величини невідомо, слід визначити вибіркову середню, а незміщену оцінку дисперсії обчислити за формулою

$$S^2=\frac{n}{n-1}\sigma^2_{выб} = \frac{\sum(x_i-\tilde{X})^2}{n-1} = \frac{(2781-2809)^2+(2836-2809)^2+(2807-2809)^2+(2763-2809)^2+(2858-2809)^2}{5-1} =$$

$$=1508{,}5\,.$$

Задача 6.2

Зроблені вимірювання випадкової величини Y при різних значеннях випадкової величини X. Визначити вибірковий коефіцієнт кореляції цих величин.

x_i	-8	-10	22	2
y_i	-10	-2	4	-1

Розв'язання

Коефіцієнт кореляції визначимо за формулою

$$r_{xy}=\frac{K_{xy}}{\sigma_x\sigma_y}.$$

Для його обчислення необхідно знайти вибірковий кореляційний момент

$$K_{xyвув}=\frac{\sum_i(x_i-\overline{X})(y_i-\overline{Y})}{n-1}.$$

Для оцінки середніх значень X та Y визначимо вибіркові середні

$$\tilde{X}=\frac{-8+10+22+2}{4}=6{,}5$$

$$\tilde{Y}=\frac{-10-2+4-1}{4}=-2{,}25.$$

Визначимо вибіркові дисперсії X та Y

$$S_x^2=\frac{(-8-6{,}5)^2+(10-6{,}5)^2+(22-6{,}5)^2+(2-6{,}5)^2}{4-1}=161;$$

$$S_y^2=\frac{(-10+2{,}25)^2+(-2+2{,}25)^2+(4+2{,}25)^2+(-1+2{,}25)^2}{4-1}=33{,}6.$$

Визначимо вибіркові середні квадратичні відхилення:

$$\sigma_{выб_x}=\sqrt{S_x^2}=\sqrt{161}=12{,}7 \qquad \sigma_{выб_y}=\sqrt{S_y^2}=\sqrt{33{,}6}=5{,}8.$$

Тепер розрахуємо кореляційний момент

$$K_{xyвув}=\frac{(-8-6{,}5)(-10+2{,}25)+(10-6{,}5)(-2+2{,}25)+(22-6{,}5)(4+2{,}25)+(2-6{,}5)(-1+2{,}25)}{4-1}=68{,}2$$

та коефіцієнт кореляції

$$r_{xy}=\frac{K_{xyвув}}{\sigma_{выб_x}\sigma_{выб_y}}=\frac{68{,}2}{12{,}7*5{,}8}=0{,}926.$$

Задача 6.3

Знайдемо довірчий інтервал для вибіркової середньої досліджуваної ознаки X. Нехай зроблено n незалежних дослідів та визначені спроможні і незміщені оцінки параметрів цієї ознаки $\tilde{X}$ та $\sigma^2_{виб}$.

Розв'язання

Нехай $\tilde{X}$ = 10, $\sigma^2_{виб}$ = 4, n = 40. Задамося значенням довірчої імовірності β = 0,95. Тоді можна записати

$$P\{(\tilde{X}-l)<\overline{X}<(\tilde{X}+l)\}=0{,}95.$$

Скористуємося тим, що випадкова величина $\tilde{X}$ є функцією n незалежних випадкових величин x_i. Тоді відповідно до центральної граничної теореми щільність розподілу випадкової величини $\tilde{X}$ практично буде підпорядковуватися нормальному закону розподілу з параметрами

$$M[\tilde{X}] = \overline{X} = 10,\ D[\tilde{X}] = \frac{\sigma^2_{ген}}{n} = 4/40 = 0,1.$$

Для нормального закону розподілу імовірність влучення випадкової величини на інтервал значень можна виразити за допомогою інтеграла імовірностей:

$$P\{(m_x^* - l) \le m_x \le (m_x^* + l)\} = \left[\Phi\left(\frac{m_x^* + l - m_x}{\sqrt{\ }\sigma_x}\right) - \Phi\left(\frac{m_x^* - l - m_x}{\sqrt{\ }\sigma_x}\right)\right] =$$

$$= \left[\Phi\left(\frac{l}{\sqrt{\ }\sigma_x}\right) - \Phi\left(\frac{-l}{\sqrt{\ }\sigma_x}\right)\right] = 2\Phi\left(\frac{l}{\sqrt{\ }\sigma_x}\right) = \beta.$$

Підставимо значення:

$$\Phi\left(\frac{l}{\sqrt{0,32}}\right) = 0,475; \qquad \frac{l}{0,453} = 1,4,$$

звідки l = 1,4 * 0,453 = 0,634.

Отже, з імовірністю 0,95 інтервал (9,366; 10,634) накриє генеральну середню досліджуваної ознаки X.

Задача 6.4

Результати вимірювання значень випадкових величин X та Y наведені у таблиці:

№ досліду	x_i	y_i
1	4	0,041
2	8	0,05
3	10	0,081
4	14	0,104
5	16	0,12
6	20	0,139
7	19	0,154
8	23	0,18
9	26	0,208
10	30	0,241
11	31	0,25
12	36	0,269
13	37	0,301

Визначити точкові й інтервальні оцінки числових характеристик системи випадкових величин X та Y, а також імовірність того, що вибіркова середня випадкової величини X відрізняється від її генеральної середньої не більш чим на 1.

Розв'язання

Проміжні розрахунки будемо зводити у таблицю. Спочатку визначимо точкові значення вибіркових середніх:

$$\tilde{Y} = \frac{\sum y_i}{n} = \frac{2{,}138}{13} = 0{,}164 \, ;$$

$$\tilde{X} = \frac{\sum x_i}{n} = \frac{274}{13} = 21{,}1 \, .$$

№ досліду	x_i	y_i	$\overset{o}{x}_i$	$\overset{o}{x}^2$	$\overset{o}{y}_i$	$\overset{o}{y}^2$	$\overset{o}{x}_i \overset{o}{y}_i$
1	4	0,041	-17,1	292,41	-0,123	0,0151	2,103
2	8	0,05	-13,1	171,61	-0,114	0,0130	1,493
3	10	0,081	-11,1	123,21	-0,083	0,0069	0,921
4	14	0,104	-7,1	50,41	-0,06	0,0036	0,426
5	16	0,12	-5,1	26,01	-0,044	0,0019	0,224
6	20	0,139	-1,1	1,21	-0,025	0,0006	0,028
7	19	0,154	-2,1	4,41	-0,01	0,0001	0,021
8	23	0,18	1,9	3,61	0,016	0,0003	0,030
9	26	0,208	4,9	24,01	0,044	0,0019	0,216
10	30	0,241	8,9	79,21	0,077	0,0059	0,685
11	31	0,25	9,9	98,01	0,086	0,0074	0,851
12	36	0,269	14,9	222,01	0,105	0,0110	1,565
13	37	0,301	15,9	252,81	0,137	0,0188	2,178
Сума	Σx_i=274	Σy_i=2,138		$\sum \overset{o}{x}{}_i^2$		$\sum \overset{o}{y}{}_i^2$	$\sum \overset{o}{x}_i \overset{o}{y}_i$

Знайдемо точкові значення незміщених вибіркових дисперсій:

$$S_y^2 = \frac{\sum_1^{13} (y_i - \tilde{Y})^2}{13-1} = \frac{\sum_1^{13} \overset{o}{y}{}^2}{12} = \frac{0{,}0866}{12} = 0{,}0072 \, ;$$

$$S_x^2 = \frac{\sum_1^{13} (x_i - \tilde{X})^2}{13-1} = \frac{\sum_1^{13} \overset{o}{x}{}^2}{12} = \frac{1348{,}93}{12} = 112{,}41 \, .$$

Визначимо вибіркові середні квадратичні відхилення

$$\sigma_{\text{виб}_\text{у}} = \sqrt{0{,}0072} = 0{,}085 \, ; \quad \sigma_{\text{виб}_\text{х}} = \sqrt{112{,}41} = 10{,}6 \, .$$

Визначимо вибірковий кореляційний момент за формулою

$$\text{K}_{\text{хувув}} = \frac{\sum_{1}^{13} \overset{o}{x}\overset{o}{y}}{13-1} = \frac{10,742}{12} = 0,895$$

та коефіцієнт кореляції за формулою

$$\text{r}_{\text{хувув}} = \frac{K_{хувув}}{\sigma_{виб_y} * \sigma_{виб_x}} = \frac{0,895}{0,085*10,6} = 0,994 \; .$$

Для визначення імовірності того, що помилка від заміни генеральної середньої випадкової величини X її вибірковою оцінкою не перевершить 1, побудуємо довірчий інтервал з межами 21,1±1 та визначимо імовірність того, що цей інтервал накриє генеральну середню випадкової величини X (довірчу імовірність β).

Скористуємося тим, що величина $\tilde{X}$ являє собою суму n=13 незалежних однаково розподілених випадкових величин x_i і, відповідно до центральної граничної теореми, при досить великому n її закон розподілу близький до нормального (а в нашім випадку ми провадили вимірювання, що завжди дає помилку, розподілену за нормальним законом). Будемо виходити з того, що величина $\tilde{X}$ розподілена за нормальним законом з характеристиками:

$$M[\tilde{X}] = \overline{X} = 21{,}1 \quad \text{та} \quad \sigma[\tilde{X}] = \sqrt{\frac{\sigma^2_{ген}}{n}} = \frac{10{,}6}{\sqrt{13}} = 2{,}94 .$$

Виразимо шукану імовірність β за допомогою функції Лапласа:

$$P\left\{\left|\tilde{X} - \overline{X}\right| < l\right\} = \beta = 2\Phi\left(\frac{l}{\sigma[\tilde{X}]}\right),$$

звідки дістанемо

$$2\Phi\left(\frac{1}{2{,}94}\right) = \beta; \; \beta = 2\Phi(0{,}34) = 0{,}266.$$

Отримане значення імовірності дуже мале, тобто подія, що полягає в тому, що помилка від заміни генеральної середньої X її вибірковою оцінкою не перевершить 1, практично неможлива. Задамося довірчою ймовірністю β =0,95 та визначимо межі відповідного їй довірчого інтервалу. Для цього запишемо вираз імовірності β

$$2\Phi\left(\frac{l}{2{,}94}\right) = \beta = 0{,}95; \; \Phi\left(\frac{l}{2{,}94}\right) = 0{,}95/2 = 0{,}475 ,$$

звідки, користуючись таблицею значень інтеграла імовірностей, дістанемо

$$\frac{l}{2{,}94} = 1{,}96 ,$$

отже, l=1,96*2,94 = 5,76. З імовірністю 0,95 інтервал (15,34; 26,86) накриє генеральну середню випадкової величини X. Помилка становить 27,3%, тобто занадто велика. Для зменшення помилки необхідно збільшити кількість вимірювань.

Побудуємо довірчий інтервал для вибіркової дисперсії випадкової величини X. Закон розподілу вибіркової дисперсії також наближається до нормального. Один з параметрів розподілу - математичне сподівання вибіркової дисперсії

$$M[S^2_x] = \sigma^2_{ген.}$$

Дисперсію вибіркової дисперсії S^2_x обчислимо за формулою:

$$D[S_x^2] = \frac{1}{n}\mu_4 - \frac{n-3}{n(n-1)}\sigma^2_{ген},$$

де μ_4 – центральний момент випадкової величини X четвертого порядку, його оцінку обчислимо за формулою:

$$\mu_4^* = \frac{1}{n}\sum_i (x_i - \tilde{X})^4.$$

Але справа в тому, що при обмеженій кількості дослідів моменти високого порядку визначаються з великими помилками, тому якщо величина X розподілена за нормальним законом, то μ_4 можна обчислити через дисперсію:

$$\mu_4 = 3 * \sigma^2_{ген}.$$

Отримаємо:

$$D[S_x^2] = \frac{3}{n}\sigma^2_{ген} - \frac{n-3}{n(n-1)}\sigma^2_{ген} = \frac{3}{13}(114{,}041)^2 - \frac{13-3}{13(13-1)}(114{,}041)^2 = 2106,$$

звідси

$$\sigma[S_x^2] = \sqrt{2106} = 45{,}9.$$

Визначимо довірчий інтервал для вибіркової дисперсії, задавшись імовірністю $\beta = 0{,}9$.

Користуючись тим, що S^2_x розподілена нормально, виразимо імовірність β за допомогою функції Лапласа:

$$P\left\{\left|S_x^2 - D_x\right| < l\right\} = \beta = 2\Phi\left(\frac{l}{\sigma[S_x^2]}\right), \text{ звідки дістанемо } \Phi\left(\frac{l}{45{,}9}\right) = 0{,}45,$$

$$\frac{l}{45{,}9} = 1{,}64, \qquad l=1{,}64*45{,}9 = 75{,}26.$$

Отже, з імовірністю 0,9 дисперсія X лежить в інтервалі 112,41±75,26. Помилка становить 67% - дуже велика.

Задача 6.5

У попередньому прикладі 6.4 результати обчислення показали, що для зменшення довірчого інтервалу (15,34; 26,86), який з довірчою імовірністю 0,95 накриє генеральну середню випадкової величини X, необхідно збільшити кількість вимірювань. Визначимо, якою повинна бути кількість вимірювань n, щоб помилка від заміни генеральної середньої випадкової величини X її вибірковою оцінкою $\tilde{X} = 21{,}1$ не перевершувала 1 з імовірністю β=0,95.

Розв'язання

Виразимо довірчу імовірність β за допомогою функції Лапласа:

$$P\left\{\left|\tilde{X}-\overline{X}\right|<l\right\}=\beta=2\Phi\left(\frac{l}{\sigma[\tilde{X}]}\right),$$

звідки дістанемо, підставивши l=1 та β =0,95,

$$\Phi\left(\frac{1}{\sigma[\tilde{X}]}\right)=0{,}475,$$

тоді $\frac{1}{\sigma[\tilde{X}]}=1{,}96$, звідки $\sigma[\tilde{X}]=\frac{1}{1{,}96}=0{,}510$. Відомо, що середнє квадратичне відхилення вибіркової середньої $\sigma[\tilde{X}]=\sqrt{\frac{\sigma^2_{ген}}{n}}$. Визначимо необхідне число вимірів n у такий спосіб:

$$n=\left(\frac{\sigma^2_{ген}}{\sigma[\tilde{X}]}\right)^2=\left(\frac{10{,}6^2}{0{,}510}\right)^2=432.$$

Отже, для досягнення необхідної точності вибіркової середньої замість 13 дослідів необхідно зробити 432 вимірювання випадкової величини X.

Практичне заняття 7
РЕГРЕСІЙНО- КОРЕЛЯЦІЙНИЙ АНАЛІЗ. МЕТОД НАЙМЕНШИХ КВАДРАТІВ

Мета — сформувати вміння на підставі статистичних даних будувати рівняння регресії та визначати його параметри за методом найменших квадратів, а також оцінювати ступінь лінійного та нелінійного зв’язку між залежною та факторною ознаками.

Кореляційний аналіз заснований на використанні рівняння регресії.

Регресією Y на X називають умовне математичне сподівання випадкової величини Y за умови, що X прийняла значення x_i. Лінію, що з'єднує точки $\overline{y}_i$, називають *лінією регресії*.

Для апроксимації лінії регресії аналітичним виразом використовують *рівняння регресії* $y_x = \varphi(\overline{x})$. Вибір виду залежності $\overline{y}_x = \varphi(x)$ можна здійснити з теоретичних міркувань або графічно за допомогою *поля кореляції***.** На практиці найчастіше використовують лінійне рівняння регресії: $Y = \rho_{yx} X + b$. Коефіцієнт при перемінній X ρ_{yx} називають *коефіцієнтом регресії*.

Для визначення параметрів залежності, що згладжує $\overline{y}_x = \varphi(x)$, і зокрема, значень параметрів ρ_{yx} та b рівняння регресії, застосовують *метод найменших квадратів* (МНК). Цей метод дозволяє при відомому класі апроксимуючої залежності $\overline{y}_x = \varphi(x)$ так вибрати значення її параметрів, щоб ця залежність щонайкраще відображала дані спостережень. При використанні МНК вимога найкращого узгодження апроксимуючої кривої $\overline{y}_x = \varphi(x)$ із дослідними даними зводиться до того, щоб сума квадратів відхилень цієї кривої від експериментальних точок оберталася у мінімум: $\sum_{i=1}^{n}(y_i - y_{ip})^2 \to min$, де y_i – значення результативної ознаки Y, отримані в результаті спостережень; y_{ip} - розрахункові значення результативної ознаки Y, отримані на підставі аналітичного вираження кривої, яка згладжує $\overline{y}_x = \varphi(x)$.

Отримані для лінійного рівняння регресії вирази для коефіцієнта регресії ρ_{yx} та вільного члена b мають вигляд:

$$\rho_{yx} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}; \quad b = \frac{\sum x_i^2 * \sum y_i - \sum x_i * \sum x_i y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}.$$

Вибірковий коефіцієнт кореляції слугує для оцінки тісноти лінійної кореляційної залежності, він визначається за формулою $r_B = \rho * \frac{\sigma_x}{\sigma_y}$.

Коефіцієнт кореляції $r_в$ дозволяє оцінити величину лінійного зв'язку між двома випадковими величинами X та Y. Вибірковий коефіцієнт кореляції приймає значення від -1 до +1 і характеризує тісноту лінійного зв'язку між ознаками у вибірці. Якщо $r_в = 0$, то лінійний зв'язок відсутній, чим ближче значення $|r_в|$ до одиниці, тим тісніше зв'язок, при $|r_в| = 1$ він стає функціональним.

Для оцінки тісноти нелінійного кореляційного зв'язку застосовують *вибіркове кореляційне відношення* η. Вибірковим кореляційним відношенням Y до X називають відношення міжгрупового середнього квадратичного відхилення до загального середнього квадратичного відхилення результативної ознаки Y: $\eta_{yx} = \frac{\sigma_{\overline{y}_x}}{\sigma_y}$, де $\sigma_{\overline{y}_x}$ -міжгрупове середнє квадратичне відхилення. Його визначають за формулою $\sigma_{\overline{y}_x} = \sqrt{D_{межгр}} = \sqrt{\frac{\sum N_j(\overline{y_j} - \overline{y})^2}{n}}$, де $\overline{y_j}$ - умовна середня значень Y j-ї групи; N_j – обсяг j-ї групи.

Міжгрупова дисперсія – це дисперсія групових середніх відносно загальної середньої.

Внутрішньогрупова дисперсія є середнім арифметичним групових дисперсій: $D_{вн.гр} = \frac{\sum N_j S_y}{n}$.

Загальна дисперсія результативної ознаки Y є сумою внутрішньогрупової та міжгрупової дисперсій: $D_y = D_{вн.гр} + D_{міжгр}$.

При функціональному зв'язку між X та Y кореляційне відношення дорівнює одиниці: $\eta_{yx} = \frac{\sigma_{\overline{y}_x}}{\sigma_y} = 1$. Якщо кореляційне відношення дорівнює нулю, η=0, то між X та Y зв'язок відсутній. Значення кореляційного відношення лежать у межах від 0 до 1: $0 \le \eta \le 1$; значення кореляційного відношення перевершують або дорівнюють вибірковому коефіцієнту кореляції: $\eta \ge |r_в|$; якщо кореляційне відношення дорівнює вибірковому коефіцієнту кореляції, $\eta = |r_в|$, то між X та Y є лінійна кореляційна залежність.

Задача 7.1

Нехай у результаті досліду отримані наступні експериментальні дані:

x_i	1	2	3
y_i	1	3	4

Потрібно визначити параметри лінійної та квадратичної залежностей для X та Y.

Розв'язання

Нанесемо на координатну площину точки з координатами (x_i,y_i) (рис. 7.1).

Із графіка видно, що точки не лежать на одній прямій, і що із зростанням X У має тенденцію до зростання.

а) нехай шукана залежність - лінійна:

$$y = a_0 + a_1x,$$

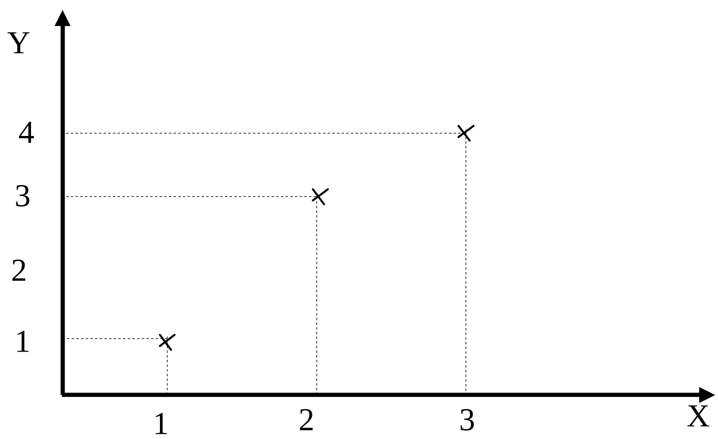


Рис. 7.1 - Побудова поля кореляції

запишемо її як функцію параметрів: $\sum_{i=1}^{n}(y_i - a_0 - a_1x_i)^2 \to min$ та візьмемо часткові похідні за параметрами a_1 та a_0 і дорівняємо їх до нуля:

$$\begin{cases} 2*\Sigma(y_i - a_0 - a_1x_i)*(-1) = 0; \\ 2*\Sigma(y_i - a_0 - a_1x_i)*(-x_i) = 0; \end{cases}$$

виконаємо перетворення:

$$\begin{cases} \Sigma a_0 + \Sigma a_1x_i = \Sigma y_i \\ \Sigma a_0\, x_i + \Sigma a_1x_i^2 = \Sigma y_i\, x_i \end{cases}$$

$$\begin{cases} na_0 + a_1\Sigma x_i = \Sigma y_i \\ a_0\Sigma\, x_i + a_1\Sigma x_i^2 = \Sigma y_i\, x_i \end{cases}$$

Підставимо значення: n=3; Σx_i =6; Σx_i^2 =14; Σy_i = 8; $\Sigma y_i\, x_i$ = 19 та визначимо параметри.

$$\begin{cases} 3a_0 + 6\,a_1 = 8 \\ 6a_0 + 14a_1 = 19. \end{cases}$$

$$\begin{cases} a_0 = -1/3 \\ a_1 = 1{,}5. \end{cases}$$

Шукана залежність має вигляд:

$$y = -1/3 + 1{,}5\ x.$$

Отримана лінійна залежність є найімовірнішою з лінійних залежностей.

Протабулюємо її та побудуємо графік (рис. 7.2).

x_i	1	2	3
y_i	1	3	4
$y_i^т$	1,17	2,67	4,17

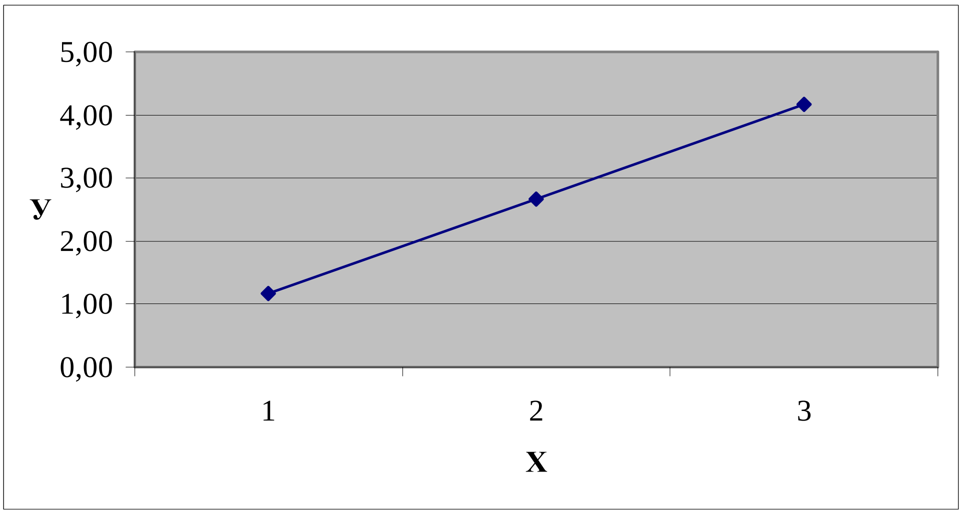


Рис. 7.2 - Лінійна залежність y = -1/3 + 1,5 x.

б) нехай шукана залежність – квадратична: $Y = a_0 + a_1x + a_2x^2$, запишемо її: $\sum_{i=1}^{n}(y_i - a_0 - a_1x_i - a_2x_i^2)^2 \rightarrow min$, візьмемо часткові похідні та дорівняємо їх до нуля:

$$\begin{cases} 2*\Sigma(y_i - a_0 - a_1x_i - a_2x_i^2)*(-1) = 0 \\ 2*\Sigma(y_i - a_0 - a_1x_i - a_2x_i^2)*(-x_i) = 0 \\ 2*\Sigma(y_i - a_0 - a_1x_i - a_2x_i^2)*(-x_i^2) = 0 \end{cases}$$

виконаємо перетворення:

$$\begin{cases} \Sigma y_i - \Sigma a_0 - \Sigma a_1 x_i - \Sigma a_2 x_i^2 = 0 \\ \Sigma y_i\, x_i - \Sigma\, a_0\, x_i - \Sigma a_1 x_i^2 - \Sigma a_2 x_i^3 = 0 \\ \Sigma y_i\, x_i^2 - \Sigma a_0\, x_i^2 - \Sigma a_1 x_i^3 - \Sigma a_2 x_i^4 = 0 \end{cases}$$

$$\begin{cases} n a_0 + a_1 \Sigma x_i + a_2 \Sigma x_i^2 = \Sigma y_i \\ a_0\, \Sigma x_i + a_1 \Sigma x_i^2 + a_2 \Sigma x_i^3 = \Sigma y_i\, x_i \\ a_0\, \Sigma x_i^2 + a_1 \Sigma x_i^3 + a_2 \Sigma x_i^4 = \Sigma y_i\, x_i^2 \end{cases}$$

Підставимо значення: n=3; Σx_i =6; Σx_i^2 =14; Σy_i = 8; $\Sigma y_i\, x_i$ = 19;Σx_i^3 =36; Σx_i^4 =98; $\Sigma y_i\, x_i^2$ = 49 і визначимо параметри

$$\begin{cases} 3a_0 + 6\, a_1 + 14 a_2 = 8 \\ 6a_0 + 14 a_1 + 36 a_2 = 19 \\ 14 a_0 + 36 a_1 + 98 a_2 = 49. \end{cases}$$

Звідки : a_2=-0,091; a_1=1,273; a_0=0,0455.

Отже, найімовірніша квадратична залежність матиме вигляд:

$$y = 0{,}0455 + 1{,}273x - 0{,}091x^2.$$

Протабулюємо її, збільшивши для наочності число точок:

x_i	0	1	2	3	4	5
y_i^T	0,05	1,23	2,23	3,05	3,68	4,14

Отримаємо графік, показаний на рис. 7.3.

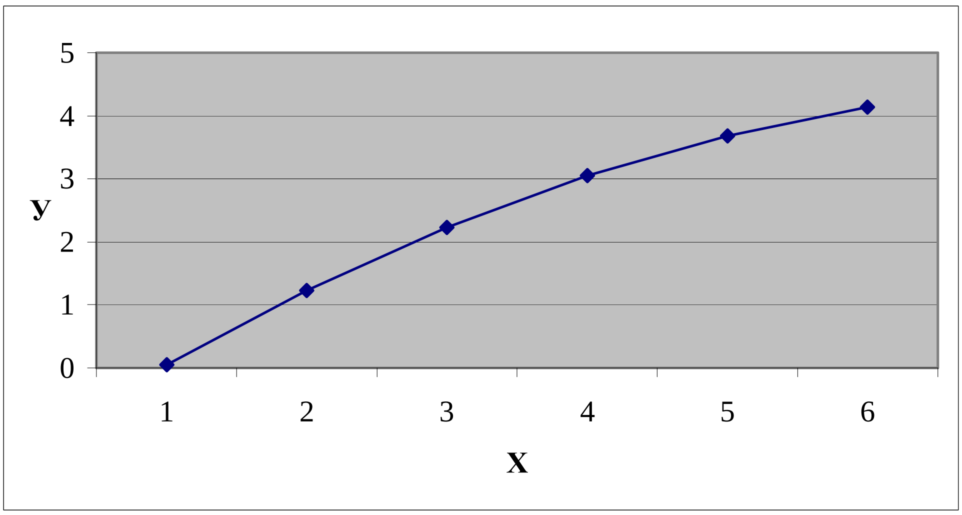


Рис. 7.3 - Квадратична залежність y = 0,0455 + 1,273x- 0,091x^2

Задача 7.2

Визначити параметри лінійної регресії для системи випадкових величин X і Y, результати вимірювання яких приведені у задачі 7.1.

№ досліду	x_i	y_i
1	4	0,041
2	8	0,05
3	10	0,081
4	14	0,104
5	16	0,12
6	20	0,139
7	19	0,154
8	23	0,18
9	26	0,208
10	30	0,241
11	31	0,25
12	36	0,269
13	37	0,301

Розв'язання

Побудуємо поле кореляції (рис. 7.4). Очевидно, що статистичні дані добре згладжуються лінійною залежністю.

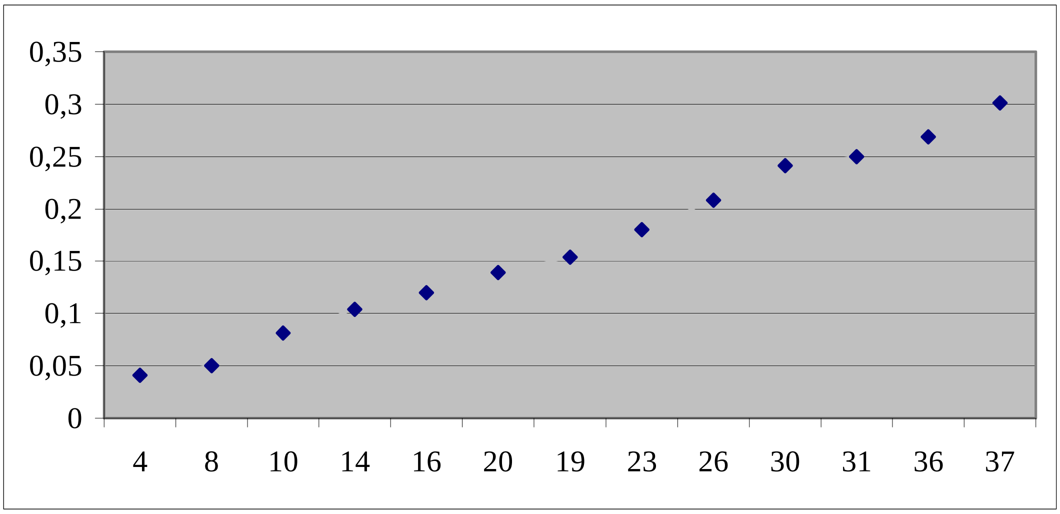


Рис. 7.4 - Побудова поля кореляції

Визначимо параметри лінійної залежності між X та Y, проміжні обчислення зробимо у таблиці

№ досліду	x_i	y_i	x_i^2	$x_i y_i$
1	4	0,041	16	0,164
2	8	0,05	64	0,4
3	10	0,081	100	0,81
4	14	0,104	196	1,456
5	16	0,12	256	1,92
6	20	0,139	400	2,78
7	19	0,154	361	2,926
8	23	0,18	529	4,14
9	26	0,208	676	5,408
10	30	0,241	900	7,23
11	31	0,25	961	7,75
12	36	0,269	1296	9,684
13	37	0,301	1369	11,137
Сума	$\Sigma x_i=274$	$\Sigma y_i=2{,}138$	$\Sigma x_i^2= 7124$	$\Sigma x_i * y_i= 55{,}805$

Отримаємо

$$\rho_{yx} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{13*55{,}08 - 274*2{,}138}{13*7124 - 274^2} = 0{,}00796\,;$$

$$b = \frac{\sum x_i^2 * \sum y_i - \sum x_i * \sum x_i y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{7124*2{,}138 - 274*55{,}8}{13*7124 - 274^2} = -0{,}0034\,.$$

Отже, шукана залежність матиме вигляд:

$$y = 0{,}00796x - 0{,}0034.$$

Задача 7.3

Визначимо значення коефіцієнта кореляції системи випадкових величин X та Y з задачі 6.4 та задачі 7.2.

Розв'язання

У результаті розв’язання попередніх задач для системи випадкових величин X та Y отримано лінійну кореляційну залежність. Скористаємося параметрами цієї залежності та визначимо коефіцієнт кореляції, розрахувавши попередньо середні квадратичні відхилення σ_x та σ_y:

$$\sigma_x = 10{,}6;\ \sigma_y = 0{,}085;$$

$$r_{вxy} = \rho * \frac{\sigma_x}{\sigma_y} = 0{,}00796 \frac{10{,}6}{0{,}085} = 0{,}994\,.$$

Задача 7.4

Визначити значення коефіцієнта кореляції системи випадкових величин I та U. Кореляційна залежність, що отримана за результатами 20 дослідів має вигляд: u = 1,525i + 3,948, середні квадратичні відхилення σ_i та σ_u.відповідно дорівнюють σ_i = 12,72; σ_u = 23,64.

Розв'язання

Розрахуємо і визначимо коефіцієнт кореляції за формулою:

$$r_{віu} = \rho * \frac{\sigma_i}{\sigma_u} = 1{,}525 \frac{12{,}72}{23{,}64} = 0{,}821\,.$$

Отже, між I та U існує тісний лінійний зв'язок.

СПИСОК ДЖЕРЕЛ

1. Гмурман В. Э. Теория вероятностей и математическая статистика. -М.: Высш. школа, 1977. - 498 с.
2. Гмурман В. Э. Руководство к решению задач по теории вероятностей и математической статистике. - М.: Высш. школа, 1975. - 330с.
3. Ачкасов А. Є., Плакіда В. Т., Воронков О. О., Воронкова Т. Б. Теорія імовірностей і математична статистика: Навчальний посібник.- Харків, ХНАМГ, 2008.- 249 с.
4. Вентцель Е. С. Теория вероятностей. - М.: Высшая школа, 1999.
5. 14Теория статистики с основами теории вероятностей: Учебное пособие для вузов/ Под ред. И. И. Елисеевой.- М.: ЮНИТИ-ДАНА, 2001.- 446 с.
6. Карасев А. И., Аксютина З. И., Савельева Т. И. Курс высшей математики для экономических вузов. ч.2. - М.:Высш. школа, 1982. - 320с.
7. Справочник по математике для экономистов (Под редакцией В.И.Ермакова.) - М.:Высш. школа, 1987. - 306с.
8. Горбань С. Ф., Снижко Н. В. Теория вероятностей и математическая статистика. - К.: МАУП, 1999.
9. http://teorver.ru
10. http://www.artspb.com
11. http://www.matburo.ru
12. http://stud-project.ru
13. http://www.statsoft.ru
14. http://www.alife.narod.ru
15. http://neuro.net.ua

ДОДАТКИ

Додаток 1

Таблиця значень функції $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0008	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Додаток 2.

Таблиця значень функції $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$

X	Ф(x)	X	Ф(x)	X	Ф(x)	X	Ф(x)
0,00	0,0000	0,36	0,1406	0,72	0,2642	1,08	0,3599
0,01	0,0040	0,37	0,1443	0,73	0,2673	1,09	0,3621
0,02	0,0080	0,38	0,1480	0,74	0,2703	1,10	0,3643
0,03	0,0120	0,39	0,1517	0,75	0,2734	1,11	0,3665
0,04	0,0160	0,40	0,1554	0,76	0,2764	1,12	0,3686
0,05	0,0199	0,41	0,1591	0,77	0,2794	1,13	0,3708
0,06	0,0239	0,42	0,1628	0,78	0,2823	1,14	0,3729
0,07	0,0279	0,43	0,1664	0,79	0,2852	1,15	0,3749
0,08	0,0319	0,44	0,1700	0,80	0,2818	1,16	0,3770
0,09	0,0359	0,45	0,1736	0,81	0,2910	1,17	0,3790
0,10	0,0398	0.46	0,1772	0,82	0,2939	1,18	0,3810
0,11	0,0438	0,47	0,1808	0,83	0,2967	1,19	0,3830
0,12	0,0478	0,48	0,1844	0,84	0,2995	1,20	0,3849
0,13	0,0517	0,49	0,1879	0,85	0,3023	1.21	0,3869
0,14	0,0557	0,50	0,1915	0,86	0,3051	1,22	0,3883
0,15	0,0596	0,51	0,1950	0,87	0,3078	1,23	0,3907
0,16	0,0636	0,52	0,1985	0,88	0,3106	1,24	0,3925
0,17	0,0675	0,53	0,2019	0,89	0,3133	1,25	0,3944
0,18	0,0714	0,54	0,2054	0,90	0,3159	1,26	0,3926
0,19	0,0753	0,55	0,2088	0,91	0,3186	1,27	0,3980
0,20	0,0793	0,56	0,2123	0,92	0,3212	1,28	0,3997
0,21	0,0832	0,57	0,2157	0,93	0,3238	1,29	0,4015
0,22	0,0871	0,58	0,2190	0,94	0,3264	1,30	0,4032
0,23	0,0910	0,59	0,2224	0,95	0,3289	1,31	0,4049
0,24	0,0948	0,60	0,2257	0,96	0,3315	1,32	0,4066
0,25	0,0987	0,61	0,2291	0,97	0,3340	1,33	0,4082
0,26	0,1026	0,62	0,2324	0,98	0,3365	1,34	0,4099
0.27	0,1064	0,63	0,2357	0,99	0,3389	1,35	0,4115
0,28	0,1103	0,64	0,2389	1,00	0,3413	1,36	0,4131
0,29	0,1141	0,65	0,2422	1,01	0,3438	1,37	0,4147
0,30	0,1179	0,66	0,2454	1,02	0,3461	1,38	0,4162
0,31	0,1217	0,67	0,2486	1,03	0,3485	1,39	0,4177
0,32	0,1255	0,68	0,2517	1,04	0,3508	1,40	0,4192
0,33	0,1293	0,69	0,2549	1,05	0,3531	1,41	0,4207
0,34	0,1331	0,70	0,2580	1,06	0,3554	1,42	0,4222
0,35	0,1368	0,71	0,2611	1,07	0,3577	1,43	0,4236

Продовження додатка 2

x	*Φ(x)*	*X*	*Φ(x)*	*X*	*Φ(x)*	*X*	*Φ(x)*
1,44	0,4251	1,73	0,4582	2,04	0,4793	2,62	0,4956
1,45	0,4265	1,74	0,4591	2,06	0,4803	2,64	0,4959
1,46	0,4279	1,75	0,4599	2,08	0,4812	2,66	0,4961
1,47	0,4292	1,76	0,4608	2,10	0,4821	2,68	0,4963
1,48	0,4306	1,77	0,4616	2,12	0,4830	2,70	0,4965
1,49	0,4319	1,78	0,4625	2,14	0,4838	2,72	0,4967
1,50	0,4332	1,79	0,4633	2,16	0,4846	2,74	0,4969
1,51	0,4345	1,80	0,4641	2,18	0,4854	2,76	0,4971
1,52	0,4357	1,81	0,4649	2,20	0,4861	2,78	0,4973
1,53	0,4370	1,82	0,4656	2,22	0,4868	2,80	0,4974
1,54	0,4382	1,83	0,4664	2,24	0,4875	2,82	0,4976
1,55	0,4394	1,84	0,4671	2,26	0,4881	2,84	0,4977
1,56	0,4406	1,85	0,4678	2,28	0,4887	2,86	0,4979
1,57	0,4418	1,86	0,4686	2,30	0,4893	2,88	0,4980
1,58	0,4429	1,87	0,4693	2,32	0,4898	2,90	0,4981
1,59	0,4441	1,88	0,4699	2,34	0,4904	2,92	0,4982
1,60	0,4452	1,89	0,4706	2,36	0,4909	2,94	0,4984
1,61	0,4463	1,90	0,4713	2,38	0,4913	2,96	0,4985
1,62	0,4474	1,91	0,4719	2,40	0,4918	2,98	0,4986
1,63	0,4484	1,92	0,4726	2,42	0,4922	3,00	0,49865
1,64	0,4495	1,93	0,4732	2,44	0,4927	3,20	0,49931
1,65	0,4505	1,94	0,4738	2,46	0,4931	3,40	0,49966
1,66	0,4515	1,95	0,4744	2,48	0,4934	3,60	0,499841
1,67	0,4525	1,96	0,4750	2,50	0,4938	3,80	0,499928
1,68	0,4535	1,97	0,4756	2,52	0,4941	4,00	0,499968
1,69	0,4545	1,98	0,4761	2,54	0,4945	4,50	0,499997
1,70	0,4554	1,99	0,4767	2,56	0,4948	5,00	0,499997
1,71	0,4564	2,00	0,4772	2,58	0,4951		
1,72	0,4573	2,02	0,4783	2,60	0,4953		

ЗМІСТ

ЗАГАЛЬНІ ПОЛОЖЕННЯ 3
МЕТОДИЧНІ ВКАЗІВКИ ДО САМОСТІЙНОЇ РОБОТИ 4
ЗМ 1. ТЕОРІЯ ІМОВІРНОСТЕЙ 4
Тема 1. ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ІМОВІРНОСТЕЙ 4
Тема 2. ОПЕРАЦІЇ НАД ПОДІЯМИ. ТЕОРЕМИ ТЕОРІЇ ІМОВІРНОСТЕЙ. ОСНОВНІ ФОРМУЛИ ТЕОРІЇ ІМОВІРНОСТЕЙ 5
Тема 3. ВИПАДКОВА ВЕЛИЧИНА ТА ЇЇ ЗАКОНИ РОЗПОДІЛУ .. 6
Тема 4. ЧИСЛОВІ ХАРАКТЕРИСТИКИ ВИПАДКОВОЇ ВЕЛИЧИНИ 7
Тема 5. НАЙВАЖЛИВІШІ ДЛЯ ПРАКТИКИ ЗАКОНИ РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН 8
Тема 6. СИСТЕМА ВИПАДКОВИХ ВЕЛИЧИН. ЗАКОНИ РОЗПОДІЛУ ТА ЧИСЛОВІ ХАРАКТЕРИСТИКИ СИСТЕМИ 9
Тема 7. ЗАКОН ВЕЛИКИХ ЧИСЕЛ 10
ЗМ 2. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ 11
Тема 8. ОБРОБКА СТАТИСТИЧНИХ ДАНИХ 11
Тема 9. ЕЛЕМЕНТИ ТЕОРІЇ КОРЕЛЯЦІЇ 12
Тема 10. ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ 13
ЗМ 3. ВИПАДКОВІ ПРОЦЕСИ 14
Тема 11. ЕЛЕМЕНТИ ТЕОРІЇ ВИПАДКОВИХ ПРОЦЕСІВ 14
МЕТОДИЧНІ ВКАЗІВКИ ДО ПРАКТИЧНИХ ЗАНЯТЬ…………… 16
Практичне заняття 1. ВИЗНАЧЕННЯ ІМОВІРНОСТІ ВИПАДКОВОЇ ПОДІЇ 16
Практичне заняття 2. ЗАКОНИ РОЗПОДІЛУ ТА ЧИСЛОВІ ХАРАКТЕРИСТИКИ ВИПАДКОВОЇ ВЕЛИЧИНИ 23
Практичне заняття 3. НАЙВАЖЛИВІШІ ДЛЯ ПРАКТИКИ ЗАКОНИ РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН 28

Практичне заняття 4. СИСТЕМА ВИПАДКОВИХ ВЕЛИЧИН. ЗАКОНИ РОЗПОДІЛУ ТА ЧИСЛОВІ ХАРАКТЕРИСТИКИ СИСТЕМИ 34
Практичне заняття 5. ОБРОБКА СТАТИСТИЧНИХ ДАНИХ. ПОБУДОВА ЗАКОНУ РОЗПОДІЛУ ВИПАДКОВОЇ ВЕЛИЧИНИ ЗА СТАТИСТИЧНИМИ ДАНИМИ 41
Практичне заняття 6. ВЛАСТИВОСТІ ВИБІРКОВИХ ЧИСЛОВИХ ХАРАКТЕРИСТИК 46
Практичне заняття 7. РЕГРЕСІЙНО- КОРЕЛЯЦІЙНИЙ АНАЛІЗ. МЕТОД НАЙМЕНШИХ КВАДРАТІВ 54
СПИСОК ДЖЕРЕЛ 62
ДОДАТКИ 63

НАВЧАЛЬНЕ ВИДАННЯ

МЕТОДИЧНІ ВКАЗІВКИ
до практичних занять
з курсу

ТЕОРІЯ ІМОВІРНОСТЕЙ ТА МАТЕМАТИЧНА СТАТИСТИКА

(для слухачів другої вищої освіти ФПО та ЗН спеціальностей
7.050107 - "Економіка підприємства", 7.050106 - "Облік і аудит")

Укладачі: доц. Охріменко Вячеслав Миколайович,
ст. викл. Воронкова Тетяна Борисівна,
ст. викл. Воронков Олексій Олександрович

Відповідальний за випуск: А. Є. Ачкасов

За авторською редакцією

Комп’ютерне верстання: О. А. Балашова

План 2011, поз. 603 М

Підп. до друку 26.10.2011 Формат 60х84 /16
Друк на ризографі. Ум. друк. арк. 4,0
Зам. № Тираж 50 пр.

Видавець і виготовлювач:
Харківська національна академія міського господарства,
вул. Революції, 12, Харків, 61002
Електронна адреса: rectorat@ksame.kharkov.ua
Свідоцтво суб’єкта видавничої справи:
ДК № 4064 від 12.05.2011